ON THE DIAMETER OF PERMUTATION GROUPS

HARALD A. HELFGOTT AND ÁKOS SERESS

Abstract. Given a finite group G and a set A of generators, the diameter $\mathrm{diam}(\Gamma(G,A))$ of the Cayley graph $\Gamma(G,A)$ is the smallest ℓ such that every element of G can be expressed as a word of length at most ℓ in $A \cup A^{-1}$. We are concerned with bounding $\mathrm{diam}(G) := \max_A \mathrm{diam}(\Gamma(G,A))$.

It has long been conjectured that the diameter of the symmetric group of degree n is polynomially bounded in n, but the best previously known upper bound was exponential in $\sqrt{n \log n}$. We give a quasipolynomial upper bound, namely,

$$\mathrm{diam}(G) = \exp\left(O((\log n)^4 \log\log n)\right) = \exp\left((\log\log |G|)^{O(1)}\right)$$

for $G = \mathrm{Sym}(n)$ or $G = \mathrm{Alt}(n)$, where the implied constants are absolute. This addresses a key open case of Babai's conjecture on diameters of simple groups. By standard results, our bound also implies a quasipolynomial upper bound on the diameter of all transitive permutation groups of degree n.

1. Introduction

1.1. **Groups and their diameters.** Let A be a set of generators for a group G. The (undirected) *Cayley graph* $\Gamma(G,A)$ is the graph whose set of vertices is $V = G$ and whose set of edges is $E = \{\{g, ga\} : g \in G, a \in A\}$. The *diameter* $\mathrm{diam}(\Gamma)$ of a graph $\Gamma(V,E)$ is defined by

$$\mathrm{diam}(\Gamma) = \max_{v_1,v_2 \in V} \min_{\substack{P \text{ a path} \\ \text{from } v_1 \text{ to } v_2}} \mathrm{length}(P). \tag{1.1}$$

In particular, the diameter of a Cayley graph $\Gamma(G,A)$ is the maximum, for $g \in G$, of the length ℓ of the shortest expression $g = a_1^{\varepsilon_1} a_2^{\varepsilon_2} \cdots a_\ell^{\varepsilon_\ell}$ with $a_i \in A$ and $\varepsilon_i \in \{-1,1\}$ for each $i = 1, \ldots, \ell$. We may define the diameter $\mathrm{diam}(G)$ of a finite group to be the maximal diameter of the Cayley graphs $\Gamma(G,A)$ for all generating sets A of G.

Much recent work on group diameters has been motivated by the following conjecture:

Conjecture 1. (*Babai, published as* [BS92, Conj. 1.7]) *For all finite simple groups G,*

$$\mathrm{diam}(G) \le (\log |G|)^{O(1)},$$

where the implied constant is absolute.

Here and henceforth, $|S|$ denotes the number of elements of a set S.

The first class of finite simple groups for which Conj. 1 was established was $\mathrm{PSL}_2(\mathbb{Z}/p\mathbb{Z})$ with p prime, by Helfgott [Hel08]. The paper [Hel08] initiated a period of intense activity [BG08a], [BG08b], [Din], [BGS10], [Hel11], [GHa], [Var], [BGS],

[PS], [BGT], [GHb], [SGV][1] on the diameter problem and the related problem of expansion properties of Cayley graphs.

As far as work in this vein on the diameter of finite simple groups is concerned, the best results to date are those of Pyber, Szabó [PS] and Breuillard, Green, Tao [BGT]. Their wide-ranging generalisation covers all simple groups of Lie type, but (just like [GHa]) the diameter estimates retain a strong dependence on the rank; thus, they prove Conj. 1 only for groups of bounded rank. The problem for the alternating groups remained wide open.[2]

These two issues are arguably related: product theorems (of the type $|A \cdot A \cdot A| \gg |A|^{1+\delta}$ familiar since [Hel08]) are false both in the unbounded-rank case and in the case of alternating groups, and the counterexamples described in both situations in [PPSS], [PS] are based on similar principles.

In the present paper we address the case of alternating (and symmetric) groups. We expect that some of the combinatorial difficulties we overcome will also arise in the context of linear groups of large rank.

For $G = \mathrm{Alt}(n)$, Conj. 1 stipulates that $\mathrm{diam}(\mathrm{Alt}(n)) = n^{O(1)}$; [BS92] refers to this special case of Conj. 1 as a "folklore" conjecture. Indeed, this has long been a problem of interest in computer science (see [KMS84], [McK84], [BHK+90], [BBS04], [BH05]). On a more playful level, bounds on the diameter of permutation groups are relevant to every permutation puzzle (e.g., Rubik's cube).

The best previously known upper bound on $\mathrm{diam}(G)$ for $G = \mathrm{Alt}(n)$ or $G = \mathrm{Sym}(n)$ was more than two decades old:

$$\mathrm{diam}(G) \leq \exp((1 + o(1))\sqrt{n \log n}) = \exp((1 + o(1))\sqrt{\log |G|}), \tag{1.2}$$

due to Babai and Seress [BS88]. (We write $\exp(x)$ for e^x.)

1.2. **Statement of results.** Recall that a function $f(n)$ is called *quasipolynomial* if $\log(f(n))$ is a polynomial function on $\log n$. Our main result establishes a quasipolynomial upper bound for $\mathrm{diam}(\mathrm{Alt}(n))$ and $\mathrm{diam}(\mathrm{Sym}(n))$.

Main Theorem. *Let $G = \mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$. Then*

$$\mathrm{diam}(G) \leq \exp\left(O((\log n)^4 \log\log n)\right),$$

where the implied constant is absolute.

The quasipolynomial bound extends to a much broader class of permutation groups. Recall that a permutation group G acting on a set Ω is called *transitive* if

$$\forall \alpha, \beta \in \Omega \quad \exists g \in G \text{ such that } g \text{ takes } \alpha \text{ to } \beta.$$

The size $|\Omega|$ of the permutation domain is called the *degree* of G.

Kornhauser et al [KMS84] and McKenzie [McK84] raised the question what classes of permutation groups may have polynomial diameter bound in their degree. A

[1] This list is not meant to be exhaustive.

[2] See, e.g., I. Pak's remarks (made already before [PS], [BGT]) on the relative difficulty of the work remaining to do in the linear case (to be finished "in the next 10 years") and of the problem on $\mathrm{Alt}(n)$, for which there was "much less hope" [Pak].

weaker, quasipolynomial bound for all transitive groups was formally conjectured in [BS92]:

Conjecture 2. ([BS92, Conj. 1.6]) *If G is a transitive permutation group of degree n then* $\mathrm{diam}(G) \leq \exp((\log n)^{O(1)})$.

Babai and Seress [BS92] linked Conj. 2 to the diameter of alternating groups:

Theorem 1.1. ([BS92, Thm. 1.4]) *If G is a transitive permutation group of degree n then*

$$\mathrm{diam}(G) \leq \exp\left(O(\log n)^3\right) \mathrm{diam}\left(\mathrm{Alt}(k)\right),$$

where $\mathrm{Alt}(k)$ *is the largest alternating composition factor of G.*

Combining our Main Theorem with Thm. 1.1, we immediately obtain

Corollary 1.2. *Conjecture 2 is true; indeed the diameter of any transitive permutation group G of degree n is*

$$\mathrm{diam}(G) \leq \exp\left(O((\log n)^4 \log\log n)\right).$$

We note that Thm. 1.1 is not only used to prove Cor. 1.2 – it is also an important ingredient in the proof of the Main Theorem (see Lemma 6.4).

It is well-known that, for any finite group G and any set of generators A of G, the eigenvalues $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \dots$ of the adjacency matrix of $\Gamma(G, A)$ satisfy

$$\lambda_0 - \lambda_1 \geq \frac{1}{\mathrm{diam}(\Gamma(G,A))^2}. \tag{1.3}$$

(See [DSC93, Cor. 1] or the references [Ald87], [Bab91], [Gan91], [Moh91] therein.) Because of (1.3), we obtain immediately that

$$\lambda_0 - \lambda_1 \geq \exp(-O((\log n)^4 \log\log n)),$$

with consequences on expansion and the mixing rate (see, e.g., [Lov96]).

Finally, the Main Theorem and Cor. 1.2 extend to directed graphs. Given $G = \langle A \rangle$, the *directed Cayley graph* $\vec{\Gamma}(G, A)$ is the graph with vertex set G and edge set $\{(g, ga) : g \in G, a \in A\}$. The diameter of $\vec{\Gamma}(G, A)$ is defined by (1.1), where "path" should be read as "directed path"; $\overrightarrow{\mathrm{diam}}(G)$ is the maximum of $\mathrm{diam}(\vec{\Gamma}(G, A))$ taken as A varies over all generating sets A of G. Thanks to Babai's bound $\overrightarrow{\mathrm{diam}}(G) = O\left(\mathrm{diam}(G) \cdot (\log|G|)^2\right)$ [Bab06, Cor. 2.3], valid for all groups G, we obtain immediately from Cor. 1.2 that

Corollary 1.3. *Let G be a transitive group on n elements. Then*

$$\overrightarrow{\mathrm{diam}}(G) \leq \exp(O((\log n)^4 \log\log n)).$$

1.3. **General approach.** An analogy underlies recent work on growth in groups: much[3] of basic group theory carries over when, instead of subgroups, we study sets

[3]Or at least results on subgroups that rely on *grosso modo* quantitative arguments. (Crucially, the orbit-stabiliser theorem carries over (Lem. 3.1); Sylow theory, which is quantitative but relies on (necessarily delicate) congruences, does not.) As [BBS04, Lem. 2.1] (in retrospect) and Prop. 5.2 in the present work make clear, probabilistic arguments in combinatorics can also carry over, provided that the desired probability distribution on a set can be approximated quickly by the action of a random walk.

that grow slowly ($|A \cdot A \cdot A| \leq |A|^{1+\varepsilon}$). This realisation is clearer in [Hel11] than in [Hel08], and has become current since then. (The term "approximate group" [Tao08] actually first arose in a different context, namely, the generalisation of some arguments in classical additive combinatorics to the non-abelian case. (See also [Hel08, §2.3], [SSV05, Lem. 4.2].) The analogy between subgroups and slowly growing sets was also explored in a model-theoretic setting in later work by Hrushovski [Hru].)

This analogy is more important than whether one works with approximate subgroups in Helfgott's sense ($|A \cdot A \cdot A| \leq |A|^{1+\varepsilon}$) or Tao's sense ([Tao08, Def. 3.7]; the two definitions are essentially equivalent, and we will actually work with neither. We could phrase part of our argument in terms of statements of the form $|A^k| \leq |A|^{1+\varepsilon}$, but k would sometimes be larger than n; applying the tripling lemma ([RT85], [Hel08, Lem. 2.2], [Tao08, Lem. 3.4]) to such statements would weaken them fatally.

There is another issue worth emphasising: the study of growth needs to be relative. We should not think simply in terms of a group acting on itself by multiplication – even if, in the last analysis, this is the only operation available to us. Rather, growth statements often need to be thought of in terms of the action of a group G on a set X, and the effect of this action on subsets $A \subseteq G$, $B \subseteq X$. (Here X may or may not be endowed with a structure of its own.) This was already clear in [Hel11, Prop. 3.1] and [GHb], and is crucial here: a key step will involve the action of a normaliser $N_G(H)$ on a subgroup $H \leq G$ by conjugation.

Our debt to previous work on permutation groups is manifold. It is worthwhile to point out that some of our main techniques are adaptations for sets of classification-free arguments[4] on the properties of subgroups of $\mathrm{Sym}(n)$ by Babai [Bab82], Pyber [Pyb93], Bochert [Boc89], and Liebeck [Lie83]. Of particular importance is Babai and Pyber's work on the order of 2-transitive groups [Bab82], [Pyb93].

We shall also utilise previous diameter bounds. Besides Thm 1.1, we shall use the main idea from [BS88] (see Lemma 3.20) and the following theorem by Babai, Beals, and Seress. For a permutation g of a set Ω, the *support* $\mathrm{supp}(g)$ is the subset of elements of Ω that are displaced by g.

Theorem 1.4. ([BBS04]) *For every $\varepsilon < 1/3$ there exists $K(\varepsilon)$ such that, if $G = \mathrm{Alt}(n)$ or $\mathrm{Sym}(n)$ and A is a set of generators of G containing an element $x \in A$ with $1 < |\mathrm{supp}(x)| \leq \varepsilon n$, then*

$$\mathrm{diam}(\Gamma(G, A)) \leq K(\varepsilon) n^8.$$

We will use this theorem repeatedly in §6.

We note that until recently Theorem 1.4 gave the largest known explicit class of Cayley graphs of $\mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$ that has polynomially bounded diameter. In late 2010, partly based on ideas from [BBS04], Bamberg et al [BGH$^+$] proved that if a set of generators of $\mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$ contains an element of support size at most $0.63n$ then the diameter of the Cayley graph is bounded by a polynomial of n.

[4] Cf. the role of [LP11] (esp. Thm. 4.2, Thm. 6.2), which, in order to provide alternatives to the Classification of Finite Simple Groups, did (both more and less generally) for subgroups what [Hel11, §5] did for sets, and was later translated back to sets for use in [BGT].

1.4. **Outline.** We begin with some general results on growth in groups, presented in terms of actions (§3.1- 3.3). We then show that large subsets of $\mathrm{Alt}(n)$ generate a copy of $\mathrm{Alt}(\Delta)$, $|\Delta| \gg n$, quickly (§3.4). This enables us to construct such a copy starting from a stabiliser chain with long orbits (§3.5). (A stabiliser chain is a nested sequence of pointwise stabilisers.) We also note that a group projecting onto $\mathrm{Alt}(\Delta)$ (Δ not tiny) must contain an element of small support (§3.6).

Section 4 uses random walks to show that, if we are given a set of generators A for $\mathrm{Alt}(n)$, we can construct a bounded-size set of generators that are short words on A (Prop. 4.6, Cor. 4.7).

Section 5 is of particular importance. It shows how to construct stabiliser chains with long orbits. The idea is to adapt Babai's splitting lemma [Bab82] from groups to sets.

Section 6 contains the core of the proof and, in particular, the main procedure that gets iterated (Prop. 6.5). We are given a stabiliser chain with long orbits; the procedure extends the chain, whose number of elements m is increased by at least $\gg m(\log m)/(\log n)$ new elements.

We now give a sketch of the procedure. (Some definitions are simplified.) We can assume the last set in the stabiliser chain – namely, a pointwise stabiliser $B^- = A_{(\alpha_1,\ldots,\alpha_m)}$ – generates an alternating group, as otherwise we can descend to a smaller group and use induction (Lemma 6.4, based on §3.6 and Thm. 1.1). We have a small set of generators for $\langle B^-\rangle$ (by §4). We let the setwise stabiliser $B^+ = A_{\{\alpha_1,\ldots,\alpha_m\}}$ (which is large, by Lemma 3.19) act on these generators by conjugation; by an orbit-stabiliser principle, either (a) an element of B^+ centralises $\langle B^-\rangle$ (and thus has small support, and we are done by Thm. 1.4) or (b) one of the orbits is large. In the latter case, we get many new elements in the pointwise stabiliser of $(\alpha_1, \ldots, \alpha_m)$. We proceed to organise them in a stabiliser chain (by repeated applications of §5), thereby extending the chain we started with.

1.5. **Acknowledgements.** We are deeply grateful to both Pablo Spiga and Nick Gill for stimulating discussions and for their constant help. Gordon Royle organised the first author's visit to Australia; if it were not for him, our collaboration might not have happened. Thanks are also due to László Babai, Martin Kassabov, Igor Pak, Peter Sarnak and Andrzej Żuk for their advice.

Ákos Seress was supported in part by the NSF and by ARC Grant DP1096525. Travel was supported in part by H. A. Helfgott's Philip Leverhulme prize. We benefited from the kind hospitality of the University of Western Australia and the École Normale Supérieure during our visits to each other's institutions.

2. Notation

We write $[n] = \{1, 2, \ldots, n\}$. For a set Ω, $\mathrm{Sym}(\Omega)$ and $\mathrm{Alt}(\Omega)$ are the symmetric and alternating groups acting on Ω. As is customary, we often write $\mathrm{Alt}(n)$ and $\mathrm{Sym}(n)$ for $\mathrm{Alt}([n])$ and $\mathrm{Sym}([n])$ - particularly when we are thinking of these groups as abstract groups as opposed to their actions.

We write $H \leq G$ to mean that H is a subgroup of G and $H \lhd G$ to mean that H is a normal subgroup. We say that a group S is a *section* of a group G if there

exist subgroups H and K of G with $K \lhd H$ and $H/K \cong S$. We denote the identity element of a group by e.

Let A be a subset of a group G. We write $A^{-1} = \{a^{-1} : a \in A\}$, $A^k = \{a_1 a_2 \cdots a_k : a_1, \ldots, a_k \in A\}$. In [Hel08], [Hel11], the first author wrote A_ℓ to mean $(A \cup A^{-1} \cup \{e\})^\ell$; this does not seem to have become standard, and would also not do here due to the potential confusion with alternating groups. (Recall that A_n is in common usage as a synonym for $\mathrm{Alt}(n)$.) We will often include $A = A^{-1}$, $e \in A$ explicitly in our assumptions so as to simplify notation. A set A with $A = A^{-1}$ is said to be *symmetric*.

We write $|A|$ for the number of elements of a set A. (All of our sets and groups are finite.) Given a group G and a subgroup $H \leq G$, we write $[G : H]$ for the index of H in G.

Let a group G act on a set X. As is customary in the study of permutation groups, given $g \in G$ and $\alpha \in X$, we write α^g for the image of α under the action of g. We speak of the *orbit* $\alpha^A = \{\alpha^g : g \in A\}$ of a point α under the action of a set A of permutations. Our actions are right actions by default: $(\alpha^g)^h = \alpha^{gh}$. In consequence, we also use right cosets by default, i.e., cosets Hg (and so G/H is the set of all such cosets). Clearly $|G/H| = [G : H]$.

We define the commutator $[g, h]$ by $[g, h] = g^{-1}h^{-1}gh$. Again, this choice is customary for permutation groups.

Define

$$\begin{aligned} A_\Sigma &= \{g \in A : \Sigma^g = \Sigma\}, && \text{(the \textit{setwise stabiliser})} \\ A_{(\Sigma)} &= \{g \in A : \forall \alpha \in \Sigma \, (\alpha^g = \alpha)\}. && \text{(the \textit{pointwise stabiliser})} \end{aligned}$$

Given a permutation $g \in \mathrm{Sym}(\Omega)$, we define its *support* $\mathrm{supp}(g)$ to be the set of elements of Ω moved by g: $\mathrm{supp}(g) = \{\alpha \in \Omega : \alpha^g \neq \alpha\}$. If a subset $\Delta \subseteq \Omega$ is invariant under g, i.e., Δ is a union of cycles of g, then we define $g|_\Delta \in \mathrm{Sym}(\Delta)$ as the *restriction* (natural projection) of g to Δ: the permutation $g|_\Delta$ acts on Δ as g does. If Δ is invariant under some $D \subseteq \mathrm{Sym}(\Omega)$ then $D|_\Delta = \{g|_\Delta : g \in D\}$.

A partition $\mathcal{B} = \{\Omega_1, \Omega_2, \ldots, \Omega_k\}$ of a set Ω is called a *system of imprimitivity* for a transitive group $G \leq \mathrm{Sym}(\Omega)$ if G permutes the sets Ω_i, for $1 \leq i \leq k$. A transitive group $G \leq \mathrm{Sym}(\Omega)$ is called *primitive* if there are only the two trivial systems of imprmitivity for G: the partition into one-element sets, and the partition consisting of one part $\Omega_1 = \Omega$.

We say that a graph (or a multigraph) is *regular* with *degree* or *valency* d if every vertex is connected to d others; that is, "degree" and "valency" of a vertex mean the same thing. In a directed graph, the *out-degree* of a vertex x is the number of edges starting at x while the *in-degree* is the number of edges terminating at x. A directed graph is called *strongly connected* if for any two vertices x, y, there is a directed path from x to y.

By $f(n) \ll g(n)$, $g(n) \gg f(n)$ and $f(n) = O(g(n))$ we mean one and the same thing, namely, that there are $N > 0$, $C > 0$ such that $f(n) \leq C \cdot |g(n)|$ for all $n \geq N$.

3. Preliminaries on sets, groups and growth

3.1. **Orbits and stabilisers.** The orbit-stabiliser theorem from elementary group theory carries over to sets. This is a fact whose importance to the area is difficult to overemphasise. It underlies already [Hel08] at a key point (Prop. 4.1); the action at stake there is that of a group G on itself by conjugation.

The setting for the theorem is the action of a group G on a set X. The *stabiliser* G_x of a point $x \in X$ is the set $\{g \in G : x^g = x\}$. Recall that we write x^g for $g(x)$ and x^A for the orbit $\{g(x) : g \in A\}$.

Lemma 3.1 (Orbit-stabiliser theorem for sets)**.** *Let G be a group acting on a set X. Let $x \in X$, and let $A \subseteq G$ be non-empty. Then*

$$|AA^{-1} \cap G_x| \geq \frac{|A|}{|x^A|}. \tag{3.1}$$

Moreover, for every $B \subseteq G$,

$$|AB| \geq |A \cap G_x||x^B|. \tag{3.2}$$

The usual orbit-stabiliser theorem is the special case $A = B = H$, H a subgroup of G.

Proof. By the pigeonhole principle, there exists an image $x' \in x^A$ such that the set $S = \{a \in A : x^a = x'\}$ has at least $|A|/|x^A|$ elements. For any $a, a' \in S$, $x^{a(a')^{-1}} = (x')^{(a')^{-1}} = x$. Hence

$$|AA^{-1} \cap G_x| \geq |SS^{-1}| \geq |S| \geq \frac{|A|}{|x^A|}.$$

Let $b_1, b_2, \dots, b_\ell \in B$, $\ell = |x^B|$, be elements with $x^{b_i} \neq x^{b_j}$ for $i \neq j$. Consider all products of the form ab_i, $a \in A \cap G_x$, $1 \leq i \leq \ell$. If two such products ab_i, $a'b_{i'}$ are equal, then $x^{b_i} = x^{ab_i} = x^{a'b_{i'}} = x^{b'_i}$. This implies $b_i = b_{i'}$. Since $ab_i = a'b_{i'}$, we conclude that $a = a'$. We have thus shown that all products ab_i, $a \in A \cap G_x$, $1 \leq i \leq \ell$, are in fact distinct. Hence

$$\begin{aligned} |AB| &\geq |(A \cap G_x) \cdot \{b_i : 1 \leq i \leq \ell\}| \\ &= |A \cap G_x| \cdot \ell = |A \cap G_x| \cdot |x^B|. \end{aligned}$$

□

As the following corollaries show, the relation between the size of A, on the one hand, and the size of orbits and stabilisers, on the other, implies that growth in the size of either orbits or stabilisers induces growth in the size of A itself.

Corollary 3.2. *Let G be a group acting on a set X. Let $x \in X$. Let $A \subseteq G$ be a non-empty set with $A = A^{-1}$. Then, for any $k > 0$,*

$$|A^{k+1}| \geq \frac{|A^k \cap G_x|}{|A^2 \cap G_x|}|A|. \tag{3.3}$$

Proof. By (3.2),

$$|A^{k+1}| \geq |A^k \cap G_x||x^A| \geq \frac{|A^k \cap G_x|}{|A^2 \cap G_x|}|A^2 \cap G_x||x^A|.$$

Since $|A^2 \cap G_x||x^A| \geq |A|$ (by (3.1)), we obtain (3.3). □

Corollary 3.3. *Let G be a group acting on a set X. Let $x \in X$. Let $A \subseteq G$ be a non-empty set with $A = A^{-1}$. Then, for any $k > 0$,*

$$|A^{k+2}| \geq \frac{|x^{A^k}|}{|x^A|}|A|. \tag{3.4}$$

Proof. By (3.2) and (3.1),

$$|A^{k+2}| \geq |A^2 \cap G_x||x^{A^k}| \geq \frac{|A|}{|x^A|}|x^{A^k}| = \frac{|x^{A^k}|}{|x^A|}|A|.$$

□

3.2. **Lemmas on subgroups and quotients.** We start by recapitulating some of the simple material in [Hel11, §7.1]. The first lemma guarantees that we can always find many elements of AA^{-1} in any subgroup of small enough index.

Lemma 3.4 ([Hel11, Lem. 7.2])**.** *Let G be a group and H a subgroup thereof. Let $A \subseteq G$ be a non-empty set. Then*

$$|AA^{-1} \cap H| \geq \frac{|A|}{r}, \tag{3.5}$$

where r is the number of cosets of H intersecting A. In particular,

$$|AA^{-1} \cap H| \geq \frac{|A|}{[G : H]}.$$

Proof. By the orbit-stabiliser principle (3.1) applied to the natural action of G on G/H by multiplication on the right.[5] (Set $x = He = H$.) □

The following two lemmas should be read as follows: growth in a subgroup gives growth in the group; growth in a quotient gives growth in the group.

Lemma 3.5 ([Hel11, Lem. 7.3])**.** *Let G be a group and H a subgroup thereof. Let $A \subseteq G$ be a non-empty set with $A = A^{-1}$. Then, for any $k > 0$,*

$$|A^{k+1}| \geq \frac{|A^k \cap H|}{|A^2 \cap H|}|A|. \tag{3.6}$$

Proof. By Cor. 3.2 applied to the action of G on G/H by multiplication on the right (with $x = He = H$). □

For a group G and a subgroup $H \leq G$, we define the coset map $\pi_{G/H} : G \to G/H$ that maps each $g \in G$ to the right coset Hg containing g.

Lemma 3.6. (essentially [Hel11, Lem. 7.3])
Let $A \subseteq G$ be a non-empty set with $A = A^{-1}$. Then, for any $k > 0$,

$$|A^{k+2}| \geq \frac{|\pi_{G/H}(A^k)|}{|\pi_{G/H}(A)|}|A|.$$

[5]Recall that we are following the convention that G/H is the set of right cosets Hg.

Proof. By Cor. 3.3, applied with G acting on $X := G/H$ by multiplication on the right and with $x := H$ seen as an element of G/H. □

The following lemma is a generalisation of Lemma 3.4.

Lemma 3.7. *Let G be a group, let H, K be subgroups of G with $H \leq K$, and let $A \subseteq G$ be a non-empty set. Then*

$$|\pi_{K/H}(AA^{-1} \cap K)| \geq \frac{|\pi_{G/H}(A)|}{|\pi_{G/K}(A)|} \geq \frac{|\pi_{G/H}(A)|}{[G:K]}.$$

In other words: if A intersects $r[G:H]$ cosets of H in G, then AA^{-1} intersects at least $r[G:H]/[G:K] = r[K:H]$ cosets of H in K. (As usual, all our cosets are right cosets.)

Proof. Since A intersects $|\pi_{G/H}(A)|$ cosets of H in G and $|\pi_{G/K}(A)|$ cosets of K in G, and every coset of K in G is a disjoint union of cosets of H in G, the pigeonhole principle implies that there exists a coset Kg of K such that A intersects at least $k = |\pi_{G/H}(A)|/|\pi_{G/K}(A)|$ cosets $Ha \subseteq Kg$. Let $a_1, \ldots, a_k$ be elements of A in distinct cosets of H in Kg. Then $a_i a_1^{-1} \in AA^{-1} \cap K$ for each $i = 1, \ldots, k$. Finally, note that $Ha_1a_1^{-1}, \ldots, Ha_ka_1^{-1}$ are k distinct cosets of H. □

Lastly, a result of a somewhat different nature. It is a version of Schreier's lemma (rewritten slightly as in [GHb, Lem. 2.10]).

Lemma 3.8 (Schreier)**.** *Let G be a group and H a subgroup thereof. Let $A \subseteq G$ with $A = A^{-1}$ and $e \in A$. Suppose A intersects each coset of H in G. Then $A^3 \cap H$ generates $\langle A \rangle \cap H$. Moreover, $\langle A \rangle = \langle A^3 \cap H \rangle A$.*

Proof. Let $C \subseteq A$ be a full set of right coset representatives of H. We wish to show that $\langle A \rangle = \langle A^3 \cap H \rangle C$. (This immediately implies both $\langle A \rangle = \langle A^3 \cap H \rangle A$ and $\langle A \rangle \cap H = \langle A^3 \cap H \rangle$.)

Clearly $e \in \langle A^3 \cap H \rangle C$. It is thus enough to show that, if $g = hc$, where $h \in \langle A^3 \cap H \rangle$ and $c \in C$, and $a' \in A$, then ga' still lies in $\langle A^3 \cap H \rangle C$. This is easily seen: since C is a full set of coset representatives, there is a $c' \in C$ with $c' = h'ca'$ for some $h' \in H$, and thus

$$ga' = hca' = h((h')^{-1})h'ca' = h((h')^{-1})c' \in \langle A^3 \cap H \rangle (A^3 \cap H)C = \langle A^3 \cap H \rangle C,$$

where we use the fact that $h' = c'(a')^{-1}c^{-1} \in A^3$. □

3.3. **Actions and generators.** The proofs of the next two lemmas share a rather simple idea.

Lemma 3.9. *Let G be a group acting transitively on a finite set X. Let $A \subseteq G$ with $A = A^{-1}$ and $G = \langle A \rangle$. Then, for any $x \in X$,*

$$x^{A^\ell} = X,$$

where $\ell = |X|$.

Proof. Consider the orbits $\{x\} \subseteq x^A \subseteq x^{A^2} \subseteq \cdots$. Let ℓ' be the smallest integer with $x^{A^{\ell'+1}} = x^{A^{\ell'}}$. As $x^{A^{\ell'+2}} = (x^{A^{\ell'+1}})^A = (x^{A^{\ell'}})^A = x^{A^{\ell'+1}} = x^{A^{\ell'}}$, we have $x^{A^{\ell'}} = x^{\langle A\rangle} = x^G = X$. Since

$$\{x\} \subsetneq x^A \subsetneq x^{A^2} \subsetneq \cdots \subsetneq x^{A^{\ell'}} = X,$$

we have $\ell' \leq |X|$. □

Lemma 3.10. *Let G be a group acting transitively on a finite set X. Let $A \subseteq G$ with $A = A^{-1}$ and $G = \langle A\rangle$. Then there is a subset $A' \subseteq A$, $|A'| < |X|$, such that $\langle A'\rangle$ acts transitively on X.*

Proof. Let $x \in X$. Let $A_1 = \{g\}$, where g is any element of A such that $x^g \neq x$. For each $i \geq 1$, let A_{i+1} be $A_i \cup \{g_i\}$, where g_i is an element of A such that $x^{\langle A_i \cup \{g_i\}\rangle} \supsetneq x^{\langle A_i\rangle}$. If no such element g_i exists, we can conclude that $x^{\langle A_i\rangle}$ is taken to itself by every $g_i \in A$. This implies that $x^{\langle A_i\rangle}$ is taken to itself by every product of elements of A, and thus $(x^{\langle A_i\rangle})^{\langle A\rangle} = x^{\langle A\rangle}$ equals $x^{\langle A_i\rangle}$.

Hence, we have a chain

$$\{x\} \subsetneq x^{\langle A_1\rangle} \subsetneq x^{\langle A_2\rangle} \subsetneq \cdots \subsetneq x^{\langle A_i\rangle} = x^{\langle A\rangle} = X.$$

Clearly $i \leq |X| - 1$, and so $|A_i| \leq |X| - 1$. Let $A' = A_i$. □

3.4. **Large subsets of** $\mathrm{Sym}(n)$**.** Let us first prove a result on large subgroups of $\mathrm{Sym}(n)$.

Lemma 3.11. *Let $n \geq 84$. Let $G \leq \mathrm{Sym}(n)$ be transitive, with a section isomorphic to $\mathrm{Alt}(k)$ for some $k > n/2$. Then G is either $\mathrm{Alt}(n)$ or $\mathrm{Sym}(n)$.*

Proof. Since $k \geq 5$, the group $\mathrm{Alt}(k)$ is simple. Hence some composition factor of G has a section isomorphic to $\mathrm{Alt}(k)$. Assume that G is imprimitive and let $\mathcal{B}$ be a non-trivial system of imprimitivity for G. Write $b = |\mathcal{B}|$ and $m = n/b$ and let K be the kernel of the action of G on $\mathcal{B}$. Since G/K is isomorphic to a subgroup of $\mathrm{Sym}(b)$, K is isomorphic to a subgroup of $\mathrm{Sym}(m)^b$ and $b, m < k$, we obtain that G has no section isomorphic to $\mathrm{Alt}(k)$, a contradiction. This shows that G is primitive.

From [PS80], we obtain that either $G \geq \mathrm{Alt}(n)$ or $|G| \leq 4^n$. Since $|G| \geq |\mathrm{Alt}(k)| = k!/2 \geq \lceil n/2\rceil!/2$, a direct computation shows that the latter case arises only for $n < 84$. □

Our aim for the rest of this subsection will be to show that, if $A \subset \mathrm{Sym}(n)$ is very large, then $A^{n^{O(1)}}$ contains a copy of $\mathrm{Alt}(\Delta)$, $|\Delta| > n/2$. The next lemma generalizes Bochert's theorem [Boc89], [DM96, Thm. 3.3B] to subsets. Recall that, for $g \in \mathrm{Sym}(\Omega)$, we define the *support* of g by $\mathrm{supp}(g) = \{\alpha \in \Omega : \alpha^g \neq \alpha\}$.

Lemma 3.12. *Let $n \geq 5$. Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$. If $\langle A\rangle$ is a primitive permutation group and $|A| > n!/(\lfloor n/2\rfloor!)$, then A^{n^4} is either $\mathrm{Alt}([n])$ or $\mathrm{Sym}([n])$.*

Proof. Given $A \subseteq \mathrm{Sym}([n])$ as in the statement of the lemma, let k be the smallest integer such that there exists $\Delta \subseteq [n]$ with $|\Delta| = k$ and $(A^2)_{(\Delta)} = \{e\}$. Let Δ be one such set.

Suppose that $k \leq n/2$. Then $\mathrm{Sym}([n])_{(\Delta)}$ has $n!/(n-k)! < |A|$ cosets in $\mathrm{Sym}([n])$. Thus, by the pigeonhole principle, there exist two distinct elements a and b of A in the same coset. Hence $ab^{-1} \in \mathrm{Sym}([n])_{(\Delta)}$, that is, $ab^{-1} \in (A^2)_{(\Delta)}$. This contradicts the definition of k. We conclude that $k > n/2$.

The set $\Omega = [n] \setminus \Delta$ has cardinality less than k, so by definition there exists $g \in (A^2)_{(\Omega)}$ with $g \neq e$. Let $\delta \in \Delta$ with $\delta^g \neq \delta$. As the set $\Delta \setminus \{\delta\}$ has cardinality less then k, by the definition of k, there exists $h \in (A^2)_{(\Delta\setminus\{\delta\})}$ with $h \neq e$. Then $\mathrm{supp}(h) \subset \Omega \cup \{\delta\}$. Necessarily, $\delta \in \mathrm{supp}(h)$, otherwise $(A^2)_{(\Delta)}$ contains the non-identity element h. Hence $\mathrm{supp}(g) \cap \mathrm{supp}(h) = \{\delta\}$ and so the commutator $x = [g, h]$ is a 3-cycle. Note that $[g, h] \in A^8$.

Now, since $\langle A\rangle$ is primitive and contains a 3-cycle, by Jordan's theorem [DM96, Thm. 3.3A] we obtain that $\langle A\rangle \geq \mathrm{Alt}([n])$. In particular, $\langle A\rangle$ is 3-transitive, and thus its action by conjugation on the set X of all 3-cycles is transitive. By Lemma 3.9,

$$x^{A^\ell} = X,$$

where $\ell = |X| = n(n-1)(n-2)/3$ and A^ℓ acts on x by conjugation. Thus

$$A^{n(n-1)(n-2)/3}[g, h]A^{n(n-1)(n-2)/3}$$

contains all 3-cycles in $\mathrm{Alt}([n])$.

Since any element of $\mathrm{Alt}([n])$ can be written as a product of at most $\lfloor n/2\rfloor$ 3-cycles, we obtain that A^{n^4-1} contains $\mathrm{Alt}([n])$. Also, if A contains an odd permutation, then $A^{n^4} = \mathrm{Sym}([n])$. □

What happens, however, if $\langle A\rangle$ is not transitive, let alone primitive? We shall see first that, if A is large, then $\langle A\rangle$ must have at least a large orbit. In the following two lemmas, we use the inequalities

$$\left(\frac{n}{e}\right)^n < n! < 3\sqrt{n}\left(\frac{n}{e}\right)^n, \tag{3.7}$$

valid for all $n \geq 1$.

Lemma 3.13. *Let $A \subseteq \mathrm{Sym}(n)$ with $|A| \geq d^n n!$, for some number d with $0.5 < d < 1$. If n is greater than a bound depending only on d, then $\langle A\rangle$ has an orbit of length at least dn.*

Proof. Let $k := \lfloor dn\rfloor$. Suppose that $|A| \geq d^n n!$ and that the longest orbit length of $\langle A\rangle$ is less than dn. We shall prove that n is bounded from above by a function depending only on d, from which the lemma follows. We start by claiming that $|\langle A\rangle| \leq k!(n-k)!$. To prove the claim, we define a sequence of groups of non-decreasing order, starting with $\langle A\rangle$ and ending with $\mathrm{Sym}(k) \times \mathrm{Sym}(n-k)$. First, we replace $\langle A\rangle$ by a direct product of symmetric groups, acting on the same orbits as A. Then, if the length of the longest orbit O_1 is $|O_1| < k$, then we add a point α to O_1 from some other orbit O_2 and replace $\mathrm{Sym}(O_1) \times \mathrm{Sym}(O_2)$ by $\mathrm{Sym}(O_1 \cup \{\alpha\}) \times \mathrm{Sym}(O_2 \setminus \{\alpha\})$ in our direct product. Finally, if the largest orbit length is k then we merge all other orbits into one.

Now, by (3.7), we have the following inequalities:

$$\begin{aligned} \left(\frac{k}{n}\right)^n \left(\frac{n}{e}\right)^n &< \left(\frac{k}{n}\right)^n n! \leq d^n n! \leq |A| \leq |\langle A\rangle| \leq k!(n-k)! \\ &< 9\sqrt{k(n-k)}\left(\frac{k}{e}\right)^k \left(\frac{n-k}{e}\right)^{n-k} \leq \frac{9}{4}n\frac{k^k(n-k)^{n-k}}{e^n}. \end{aligned} \tag{3.8}$$

Simplifying the left-hand side together with the right-hand side, we obtain $k^{n-k} < \frac{9}{4}n(n-k)^{n-k}$, that is, $\left(\frac{k}{n-k}\right)^{n-k} < \frac{9}{4}n$.

We define $c := \left(\frac{d}{1-d}\right)^{1-d}$. As

$$\lim_{n\to\infty}\left(\frac{k}{n-k}\right)^{\frac{n-k}{n}} = c > 1,$$

for large enough n, depending only on d, we have $\left(\frac{k}{n-k}\right)^{n-k} > \left(\frac{1+c}{2}\right)^n$. However, $\left(\frac{1+c}{2}\right)^n < \frac{9}{4}n$ is false if n is greater than a bound depending only on d, proving our claim. □

Using Bochert's theorem [Boc89], Liebeck [Lie83] derived a result on large subgroups of $\mathrm{Sym}(n)$ that does not assume transitivity or primitivity; the following lemma is a generalisation to sets. Note that we use only elementary counting arguments, whereas [Lie83] is based on a detailed examination of the subgroup structure of $\mathrm{Sym}(n)$.

Lemma 3.14. *Let d be a number with $0.5 < d < 1$. If $A \subseteq \mathrm{Sym}([n])$ (with $A = A^{-1}$) has cardinality $|A| \geq d^n n!$ and n is larger than a bound depending only on d, then there exists an orbit $\Delta \subseteq [n]$ of $\langle A\rangle$ such that $|\Delta| \geq dn$ and $(A^{n^4})|_\Delta$ is $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$.*

Proof. By Lemma 3.13, for large enough n the group $B = \langle A\rangle$ has an orbit Δ of length $k \geq dn$. Write $\rho = k/n$ and note that $d \leq \rho \leq 1$. The group $G = B|_\Delta$ has order at least $d^n n!/(n-k)!$, so estimating $k!(n-k)!$ from above as in (3.8) and estimating $n!$ from below by (3.7), we obtain

$$\begin{aligned} [\mathrm{Sym}(\Delta) : G] &\leq \frac{k!(n-k)!}{d^n n!} < \frac{9}{4}n\frac{k^k(n-k)^{n-k}}{d^n n^n} = \\ \frac{9}{4}n\left(\frac{\rho^\rho(1-\rho)^{1-\rho}}{d}\right)^n &= \frac{9}{4}n\left(2^{\frac{1}{\rho}}\rho(1-\rho)^{\frac{1-\rho}{\rho}}\right)^{\rho n}\left(\frac{1}{2d}\right)^n. \end{aligned} \tag{3.9}$$

Next, we show that for large values of n the transitive group G cannot be imprimitive. Indeed, if G is imprimitive, then using (3.7) we have

$$[\mathrm{Sym}(\Delta) : G] \geq \frac{1}{2}\binom{k}{\lfloor k/2\rfloor} > \frac{1}{2}\frac{(\frac{k}{e})^k}{9k(\frac{k}{2e})^k} > \frac{1}{18n}2^{\rho n}. \tag{3.10}$$

A direct computation shows that the function $f(\rho) = 2^{1/\rho}\rho(1-\rho)^{(1-\rho)/\rho}$ is monotone increasing in the interval $[1/2, 1)$ with supremum 2. Hence, comparing the upper

and lower bounds for $[\mathrm{Sym}(\Delta) : G]$ deduced in (3.9) and (3.10), we obtain

$$\frac{9}{4} n 2^{\rho n} \left(\frac{1}{2d}\right)^n > \frac{1}{18n} 2^{\rho n}. \tag{3.11}$$

As $d > 1/2$, for large enough n we have $(2d)^n > (18n)(\frac{9}{4}n)$ and therefore (3.11) cannot hold.

Hence G is primitive and $A|_\Delta$ is a set of size at least $d^n n!/(n-k)! \geq d^n k! > k!/(\lfloor k/2 \rfloor)!$ (where the last inequality holds for n greater than a lower bound depending only on d). Therefore, by Lemma 3.12, $(A|_\Delta)^{n^4}$ is either $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$, and hence so is $(A^{n^4})|_\Delta = (A|_\Delta)^{n^4}$. □

The following auxiliary lemma will allow us to strengthen the conclusions of Lemma 3.14. It shows that, if a set A contains a copy of Alt in a weak sense, A^{8n} contains a copy of Alt in a strong sense.

Lemma 3.15. *Let $n \geq 13$. Let $\Delta \subseteq [n]$, $|\Delta| > \frac{n}{2}$. Let $A \subseteq \mathrm{Sym}([n])_\Delta$, with $A = A^{-1}$ and $e \in A$, be such that Δ is an orbit of $\langle A \rangle$ and $A|_\Delta$ is $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$. Then $(A^{8n})_{([n]\setminus\Delta)}|_\Delta$ contains $\mathrm{Alt}(\Delta)$.*

Proof. It is clear that $|A| \geq |\mathrm{Alt}(\Delta)| > |\mathrm{Sym}([n] \setminus \Delta)|$. Thus, by the pigeonhole principle, there are $h_1, h_2 \in A$, $h_1 \neq h_2$, such that $h_1|_{[n]\setminus\Delta} = h_2|_{[n]\setminus\Delta}$, and so $g = h_1 h_2^{-1}$ fixes $[n] \setminus \Delta$ pointwise.

We show that $(A^{14})_{([n]\setminus\Delta)}$ contains an element g' such that $g'|_\Delta$ is a 3-cycle. If $g|_\Delta$ has at least two fixed points then there exists an element $h \in A$ so that $h|_\Delta$ is a 3-cycle, with $\mathrm{supp}(h|_\Delta)$ intersecting $\mathrm{supp}(g|_\Delta)$ in exactly one point. Then $g' = [g, h] \in A^{2+1+2+1} = A^6$ fixes $[n] \setminus \Delta$ pointwise and $g'|_\Delta$ is a 3-cycle. If g contains a cycle $(\alpha\beta\gamma\delta\dots)$ of length at least 4, then we choose an element $h \in A$ with $h|_\Delta = (\alpha\beta\gamma)$ and let $g' = [g, h] \in A^6$. Then g' fixes $[n] \setminus \Delta$ pointwise and $g'|_\Delta$ is the 3-cycle $(\alpha\beta\delta)$.

In all other cases, $|\mathrm{supp}(g|_\Delta)| \geq |\Delta| - 1 \geq 6$ and all nontrivial cycles of g have length 2 or 3. Hence $g|_\Delta$ contains at least two 3-cycles, or at least two 2-cycles. If $g|_\Delta$ contains the cycles $(\alpha\beta\gamma)$ and $(\delta\eta\nu)$ then we choose an element $h \in A$ with $h|_\Delta = (\alpha\eta)(\beta\delta\gamma\nu)$. A little computation shows that $g' = [g, h]$ fixes $[n] \setminus \Delta$ pointwise and $g'|_\Delta$ is the 3-cycle $(\delta\eta\nu)$.

Finally, suppose g contains the 2-cycles $(\alpha\beta)$ and $(\gamma\delta)$. We choose again an element $h \in A$ with $h|_\Delta = (\alpha\beta\gamma)$; then $\mathrm{supp}([g, h]) = \{\alpha, \beta, \gamma, \delta\}$ and $[g, h]$ fixes $[n] \setminus \Delta$ pointwise. Since $[g, h] \in A^6$ also fixes at least two points of Δ, we deduce as in the very first case of our analysis that the commutator $g' = [[g, h], h']$ with an appropriate $h' \in A$ is a 3-cycle. Note that $g' \in A^{6+1+6+1} = A^{14}$.

Given any 3-cycle s in $\mathrm{Sym}(\Delta)$, we can conjugate g' by an appropriate element of A to get an element of $(A^{16})_{([n]\setminus\Delta)}$ whose restriction to Δ equals s. Now, every element of $\mathrm{Alt}(\Delta)$ is the product of at most $\lfloor |\Delta|/2 \rfloor$ 3-cycles. Hence $(A^{16\lfloor n/2 \rfloor})_{(\lfloor n \rfloor \setminus \Delta)}|_\Delta$ contains $\mathrm{Alt}(\Delta)$. □

We finally obtain the statement we need on large sets A: a power of A (namely, A^{8n^5}) contains elements acting as the full alternating group on an orbit Δ, and leaving all points outside Δ fixed.

Proposition 3.16. *Let d be a number with $0.5 < d < 1$. Let $A \subseteq \mathrm{Sym}(n)$ with $A = A^{-1}$ and $e \in A$. If $|A| \geq d^n n!$ and n is larger than a bound depending only on d, then there exists an orbit $\Delta \subseteq [n]$ of $\langle A \rangle$ such that $|\Delta| \geq dn$ and $(A^{8n^5})_{([n]\setminus\Delta)}|_\Delta$ contains $\mathrm{Alt}(\Delta)$.*

Proof. Immediate from Lemma 3.14 and Lemma 3.15, where the latter is applied with $(A^{n^4})_\Delta$ instead of A. □

3.5. **Bases and stabiliser chains.** Given a permutation group G on a set Ω, a subset Σ of Ω is called a *base* if $G_{(\Sigma)} = \{e\}$. This definition goes back to Sims [Sim70]. If, instead of G, we consider a subset A of $\mathrm{Sym}(\Omega)$, then, as the following lemma suggests, it makes sense to see whether $(AA^{-1})_{(\Sigma)}$ (rather than $A_{(\Sigma)}$) equals $\{e\}$.

Lemma 3.17. *Let $A \subseteq \mathrm{Sym}(\Omega)$, $|\Omega| = n$. If $\Sigma \subseteq \Omega$ is such that $(AA^{-1})_{(\Sigma)} = \{e\}$, then $|\Sigma| \geq \log_n |A|$.*

Proof. Notice first that $[\mathrm{Sym}(\Omega) : (\mathrm{Sym}(\Omega))_{(\Sigma)}] \leq n^{|\Omega|}$. By the pigeonhole principle, if $|A| > n^{|\Sigma|}$, then there exists a right coset of $(\mathrm{Sym}(\Omega))_{(\Sigma)}$ containing more than one element of A, and thus

$$|(AA^{-1})_{(\Sigma)}| = |AA^{-1} \cap (\mathrm{Sym}(\Omega))_{(\Sigma)}| > 1.$$

Hence, if $(AA^{-1})_{(\Sigma)} = \{e\}$, then we have $|A| \leq n^{|\Sigma|}$, i.e., $|\Sigma| \geq \log_n |A|$. □

The use of stabiliser chains $H > H_{\alpha_1} > H_{(\alpha_1,\alpha_2)} > \cdots$ is very common in computational group theory (starting, again, with the work of Sims; see references in [Ser03, §4.1]). We may study a similar chain $A > A_{\alpha_1} > A_{(\alpha_1,\alpha_2)} > \cdots$ when A is merely a set.

Lemma 3.18. *Let $\Sigma = \{\alpha_1, \ldots, \alpha_m\} \subseteq [n]$ and $A \subseteq \mathrm{Sym}([n])$. Suppose that*

$$\left|\alpha_i^{A_{(\alpha_1,\ldots,\alpha_{i-1})}}\right| \geq r_i$$

for all $i = 1, 2, \ldots, m$. Then A^m intersects at least $\prod_{i=1}^m r_i$ cosets of $\mathrm{Sym}([n])_{(\Sigma)}$.

Proof. For each $1 \leq i \leq m$, write $\Delta_i = \alpha_i^{A_{(\alpha_1,\ldots,\alpha_{i-1})}}$; thus $|\Delta_i| \geq r_i$. For each $\delta \in \Delta_i$, pick $g_\delta \in A_{(\alpha_1,\ldots,\alpha_{i-1})}$ with $\alpha_i^{g_\delta} = \delta$ and write $S_i = \{g_\delta : \delta \in \Delta_i\}$. Clearly, $|S_i| = |\Delta_i|$ and $S_i \subseteq A$. We show that for every two distinct tuples

$$(s_1, s_2, \ldots, s_m), (s_1', s_2', \ldots, s_m') \in S_1 \times \cdots \times S_m$$

the products $P = s_m s_{m-1} \cdots s_1$ and $P' = s_m' s_{m-1}' \cdots s_1'$ belong to two distinct cosets of $\mathrm{Sym}([n])_{(\Sigma)}$. From this it follows that A^m intersects at least $|S_1| \cdots |S_m| = |\Delta_1| \cdots |\Delta_m| \geq \prod_{i=1}^m r_i$ cosets of $\mathrm{Sym}([n])_{(\Sigma)}$.

We argue by contradiction, that is, we assume that P and P' map $(\alpha_1, \ldots, \alpha_m)$ to the same m-tuple. Let j be the smallest index such that $s_j \neq s_j'$. Then $Q =$

$Ps_1^{-1}\cdots s_{j-1}^{-1}$ and $Q' = P's_1^{-1}\cdots s_{j-1}^{-1} = P's_1'^{-1}\cdots s_{j-1}'^{-1}$ also map $(\alpha_1,\ldots,\alpha_m)$ to the same m-tuple. Note that for all $k\leq m$, s_k and s_k' fix $(\alpha_1,\ldots,\alpha_{k-1})$ pointwise. Thus

$$\alpha_j^Q = \alpha_j^{s_j} \neq \alpha_j^{s_j'} = \alpha_j^{Q'},$$

contradicting our assumption. □

We thus see that, if we choose $\alpha_1,\alpha_2,\ldots$ so that the orbits $\alpha_i^{A_{(\alpha_1,\ldots,\alpha_{i-1})}}$ are large, we get to occupy many cosets of $(\mathrm{Sym}([n]))_{(\Sigma)}$. By Lemma 3.7, this will enable us to occupy many cosets of $(\mathrm{Sym}([n]))_{(\Sigma)}$ in the setwise stabiliser $(\mathrm{Sym}([n]))_\Sigma$. We will then be able to apply Prop. 3.16 to build a large alternating group within $\mathrm{Sym}(\Sigma)\cong(\mathrm{Sym}([n]))_\Sigma/(\mathrm{Sym}([n]))_{(\Sigma)}$. This procedure is already implicit in [Pyb93, Lem. 3]; indeed, what amounts to this is signalled by Pyber as the main new element in his refinement [Pyb93, Thm. A] of Babai's theorem on the order of doubly transitive groups [Bab82]. The main difference is that we have to work, of course, with sets rather than groups; we also obtain a somewhat stronger conclusion due to our using Prop. 3.16 rather than invoking Liebeck's lemma directly.

Lemma 3.19. *Let $A\subseteq\mathrm{Sym}([n])$ with $A=A^{-1}$ and $e\in A$. Let $\Sigma=\{\alpha_1,\ldots,\alpha_m\}\subseteq[n]$ be such that*

$$\left|\alpha_i^{A_{(\alpha_1,\ldots,\alpha_{i-1})}}\right| \geq dn \tag{3.12}$$

for all $i=1,2,\ldots,m$, where $d>0.5$. Then, provided that m is larger than a bound $C(d)$ depending only on d, there exists $\Delta\subseteq\Sigma$ with $|\Delta|\geq d|\Sigma|$ and

$$\mathrm{Alt}(\Delta)\subseteq((A^{16m^6})_\Sigma)_{(\Sigma\setminus\Delta)}|_\Delta.$$

Proof. By (3.12) and Lemma 3.18, A^m intersects at least $(dn)^m$ cosets of $\mathrm{Sym}([n])_{(\Sigma)}$ in $\mathrm{Sym}([n])$. Since

$$[\mathrm{Sym}([n]):\mathrm{Sym}([n])_\Sigma] = \frac{[\mathrm{Sym}([n]):\mathrm{Sym}([n])_{(\Sigma)}]}{[\mathrm{Sym}([n])_\Sigma:\mathrm{Sym}([n])_{(\Sigma)}]} \leq \frac{n^m}{m!},$$

Lemma 3.7 implies (with $G=\mathrm{Sym}([n])$, $K=\mathrm{Sym}([n])_\Sigma$, $H=\mathrm{Sym}([n])_{(\Sigma)}$, and A^m instead of A) that

$$|\pi_{K/H}(A^{2m}\cap K)| \geq \frac{|\pi_{G/H}(A^m)|}{n^m/m!} \geq \frac{(dn)^m}{n^m/m!} = d^m m!.$$

Note that $|\pi_{K/H}(A^{2m}\cap K)| = \left|(A^{2m})_\Sigma|_\Sigma\right|$. We can thus apply Prop. 3.16 (with m instead of n, and $A'=(A^{2m})_\Sigma|_\Sigma$ instead of A) and obtain that there is a set $\Delta\subseteq\Sigma$ such that $|\Delta|\geq dm$ and $((A')^{8m^5})_{(\Sigma\setminus\Delta)}|_\Delta$ contains $\mathrm{Alt}(\Delta)$. □

3.6. **Existence of elements of small support.** The following lemma is essentially [BS87, Lemma 3] (or [BS88, Lemma 1]).

Lemma 3.20. *Let $\Delta\subseteq[n]$, $|\Delta|\geq c(\log n)^2$, $c>0$. Let $H\leq(\mathrm{Sym}(n))_\Delta$. Assume $H|_\Delta$ is $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$.*

Let Γ be any orbit of H. Then, if n is larger than a bound depending only on c, H contains an element g with $g|_\Delta\neq 1$ and $|\operatorname{supp}(g|_\Gamma)|<|\Gamma|/4$.

Proof. Let $p_1 = 2$, $p_2 = 3, \ldots$, p_k be the sequence of the first k primes, where k is the least integer such that $p_1 p_2 \cdots p_k > n^4$. Much as in [BS87], we remark that, by elementary bounds towards the prime number theorem,

$$2p_1 + p_2 + \cdots + p_k < c(\log n)^2, \tag{3.13}$$

provided that n be larger than a bound depending only on c. Thus H contains an element h such that $h|_\Delta$ consists of $|\Delta| - (2p_1 + p_2 + \cdots + p_k)$ fixed points and cycles of length $p_1, p_1, p_2, p_3, \ldots, p_k$. (We need two cycles of length $p_1 = 2$ because we want an even permutation on Δ.)

We can now reason as in [BS87, Lemma 3] or [BS88, Lemma 1]. For every $\gamma \in \Gamma$, denote by κ_γ the length (possibly 1) of the cycle of h containing γ and for $i \leq k$ define $\Gamma_i := \{\gamma \in \Gamma : p_i \mid \kappa_\gamma\}$. Then

$$\sum_{\gamma \in \Gamma} \sum_{p_i | \kappa_\gamma} \log p_i < |\Gamma| \log n \tag{3.14}$$

because $\kappa_\gamma < n$ implies that for all γ the inner sum is less than $\log n$. Exchanging the order of summation,

$$\sum_{\gamma \in \Gamma} \sum_{p_i | \kappa_\gamma} \log p_i = \sum_{i=1}^{k} |\Gamma_i| \log p_i.$$

If $|\Gamma_i| \geq |\Gamma|/4$ for all $i \leq k$ then

$$\sum_{i=1}^{k} |\Gamma_i| \log p_i \geq \frac{|\Gamma|}{4} \log \left(\prod_{i=1}^{k} p_i \right) > \frac{|\Gamma|}{4} \log(n^4) = |\Gamma| \log n,$$

contradicting (3.14). Hence there is a prime $p \leq p_k$ such that $p | \kappa_\gamma$ for fewer than $|\Gamma|/4$ elements γ of Γ. Denoting the order of h by $|h|$, we define $g = h^\ell$ for $\ell := |h|/p$. We obtain that $|\operatorname{supp}(g|_\Gamma)| < |\Gamma|/4$. We also have that g is non-trivial, since $g|_\Delta$ contains a p-cycle. Clearly $g \in H$, and so we are done.[6] □

4. Random walks and generation

4.1. **Random walks.** Let Γ be a strongly connected directed multigraph with vertex set $V = V(\Gamma)$. For $x \in V(\Gamma)$, we denote by $\Gamma(x)$ the multiset of endpoints of the edges starting at x (counted with multiplicities in case of multiple edges). We are interested in the special case when Γ is *regular of valency* d (i.e., $|\Gamma(x)| = d$, for each $x \in V(\Gamma)$) and Γ is also *symmetric* in the sense that for all vertices $x, y \in V(\Gamma)$, the number of edges connecting x to y is the same as the number of edges connecting y to x. These two conditions imply that the adjacency matrix A of Γ is symmetric and all row and column sums are equal to d.

A *lazy random walk* on Γ is a stochastic process where a particle moves from vertex to vertex; if the particle is at vertex x such that $\Gamma(x) = \{y_1, \ldots, y_d\}$, then the particle

[6] Since we need only the existence of g for the moment, we are not concerned by the fact that l is very large. Compare this to the situation in [BS88], where the use of a large l causes diameter bounds much weaker than those in the present paper.

- stays at x with probability $\frac{1}{2}$;
- moves to vertex y_i with probability $\frac{1}{2d}$, for all $i = 1, \ldots, d$.

Here we are concerned with the asymptotic rate of convergence for the probability distribution of a particle in a lazy random walk on Γ. For $x, y \in V(\Gamma)$, write $p_k(x, y)$ for the probability that the particle is at vertex y after k steps of a lazy random walk starting at x. For a fixed $\varepsilon > 0$, the *mixing time for* ε is the minimum value of k such that

$$\frac{1}{|V(\Gamma)|}(1-\varepsilon) \leq p_k(x,y) \leq \frac{1}{|V(\Gamma)|}(1+\varepsilon)$$

for all $x, y \in V(\Gamma)$.

We can give a crude (and well-known; see, e.g., [BBS04, Fact 2.1]) upper bound on the mixing time for regular symmetric multigraphs in terms of $N = |V(\Gamma)|$, ε and the valency d alone.

Lemma 4.1. *Let Γ be a connected regular multigraph of valency d and with N vertices. Then the mixing time for ε is at most $N^2 d \log(N/\varepsilon)$.*

Proof. Let A be the adjacency matrix of Γ. Since A is symmetric, the eigenvalues of A are real; moreover, their modulus is clearly no more than d in magnitude. Let

$$d = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_N \geq -d$$

be the eigenvalues of A and write $P = I/2 + A/2d$, where I is the $N \times N$-identity matrix. The matrix P is the probability transition matrix for the Markov process described by a lazy random walk on Γ.

The sum of every row or column of P is 1, i.e., P is a *doubly stochastic* matrix. The eigenvalues of P are

$$1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$$

with $\lambda_i = 1/2 + \mu_i/2d$ for each $i = 1, \ldots, N$. It is well-known that the asymptotic rate of convergence to the uniform distribution of a lazy random walk is determined by λ_2 (see for example [Lov96]): for each vertex y of Γ and for each $k \geq 1$, we obtain from [Lov96, Theorem 5.1] that $|p_k(x, y) - 1/N| \leq \lambda_2^k$.

By [Fie72, Lemma 2.4 and Theorem 3.4], we have

$$\lambda_2 \leq 1 - 2(1 - \cos(\pi/N))\mu(P),$$

where $\mu(P) = \min_{\emptyset \neq M \subseteq V} \sum_{i \in M, j \notin M} p_{ij}$. As Γ is a connected regular graph of valency d, we have $\mu(P) \geq 1/2d$. Using the Taylor series for $\cos(x)$, we see that $(1 - \cos(\pi/N)) \geq 1/N^2$. Thence $|p_k(x, y) - 1/N| \leq (1 - 1/(N^2 d))^k$. Since the series $\{(1 - 1/n)^n\}_{n \geq 1}$ is monotone increasing converging to e^{-1}, for $k \geq N^2 d \log(N/\varepsilon)$, we obtain that $|p_k(x, y) - 1/N| \leq \varepsilon/N$, as desired. □

We will generally study regular symmetric multigraphs of the following type. Let G be a group and A be a subset of G with $A = A^{-1}$ and $e \in A$. Let G act on a set X. We take the elements of X as the vertices of our multigraph, and draw one edge from $x \in X$ to $x' \in X$ for every $a \in A$ such that $x^a = x'$. A walk on the graph then corresponds to the action of an element of A^ℓ on an element x of X, where ℓ is the length of the walk and x is the starting point of the walk.

Lemma 4.1 then gives us a lower bound on how large ℓ has to be for the action of A^ℓ on X to have a rather strong randomising effect. This idea was already exploited in [BBS04, §2], with $G = \mathrm{Sym}([n])$ and X the set of k-tuples of $[n]$.

Lemma 4.2. *Let H be a k-transitive subgroup of $\mathrm{Sym}([n])$. Let A be a set of generators of H with $A = A^{-1}$ and $e \in A$. Then there is a subset $A' \subseteq A$ with $A' = (A')^{-1}$, such that, for every $\varepsilon > 0$, for any $\ell \geq 2n^{3k}\log(n^k/\varepsilon)$, and for any k-tuples $\overline{x} = (x_1, \dots, x_k)$, $\overline{y} = (y_1, \dots, y_k)$ of distinct elements of $[n]$, the probability of the event*

$$\overline{y} = \overline{x}^{g_1 g_2 \cdots g_\ell}$$

for $g_1, \dots, g_\ell \in A'$ (chosen independently, with uniform distribution on $A' \setminus \{e\}$ and with the identity being assigned probability $1/2$) is at least $(1-\varepsilon)\frac{(n-k)!}{n!}$ and at most $(1+\varepsilon)\frac{(n-k)!}{n!}$.

Proof. Let Δ be the set of k-tuples of distinct elements of $[n]$. Since H acts transitively on Δ and since $\langle A \rangle = H$, Lemma 3.10 gives us a subset A' of A with $\langle A' \rangle$ transitive on Δ and with $|A'| < |\Delta|$. Set $A_0 = A' \cup A'^{-1}$. Let Γ be the multigraph with vertex set Δ and with $\Gamma(\overline{x}) = \{\overline{x}^a \mid a \in A_0\}$ as the multiset of neighbours of $\overline{x}$ for each $\overline{x} \in \Delta$. Clearly, Γ is a regular graph of valency $|A_0| \leq 2|\Delta|$ and with $|\Delta| \leq n^k$ vertices. Now the statement follows from Lemma 4.1 applied to Γ. □

4.2. **Generators.** Given $A \subseteq \mathrm{Sym}([n])$ such that $\langle A \rangle$ is $\mathrm{Alt}([n])$ or $\mathrm{Sym}([n])$, how long can it take to construct a *small* set of generators for a *transitive* subgroup of $\langle A \rangle$? This subsection is devoted to answering that question. We start with proving two auxiliary lemmas.

Lemma 4.3. *Let $A \subset \mathrm{Sym}([n])$, $e \in A$. Assume $\langle A \rangle$ is transitive. Then there is a $g \in A^n$ such that $|\mathrm{supp}(g)| \geq n/2$.*

Proof. For each $i \in [n]$, let g_i be an element of A moving i. (If no such element existed, then $\langle A \rangle$ could not be transitive.) Let $g = g_1^{r_1} g_2^{r_2} \dots g_n^{r_n}$, where $r_1, r_2, \dots, r_n \in \{0, 1\}$ are independent random variables taking the values 0 and 1 with equal probability. (Such an element g is called a *random subproduct* of the sequence (g_i), see, e.g., [Ser03, §2.3] for other applications).

Let $\alpha \in [n]$ be arbitrary. Let j be the largest integer such that g_j moves α. Then g moves α if and only if $g' = g_1^{r_1} \dots g_j^{r_j}$ moves α. Take $r_1, r_2, \dots, r_{j-1}$ as given. If $\beta = \alpha^{g_1^{r_1} \dots g_{j-1}^{r_{j-1}}}$ equals α, then g' moves α if and only if $r_j = 1$; this happens with probability $1/2$. If $\beta \neq \alpha$, then g' certainly moves α if $r_j = 0$, and thus moves α with probability at least $1/2$. Thus g moves α with probability at least $1/2$.

Summing over all α, we see that the expected value of the number of elements of $[n]$ moved by g is at least $n/2$. In particular, there is a $g \in A^n$ moving at least $n/2$ elements of $[n]$. □

The proof of the following lemma may seem familiar; its basic idea is common in sphere-packing arguments.

Lemma 4.4. *For a positive integer n, let $k = \lfloor \log_2 n \rfloor = \lfloor (\log n)/(\log 2) \rfloor$ and $U = \{0,1\}^{5k}$ the set of $0,1$-sequences of length $5k$. If n is larger than an appropriate absolute constant then there exists $V \subseteq U$, $|V| > n$ such that any two sequences in V differ in more than k coordinates.*

Proof. We may construct the required set $V = \{v_1, v_2, \ldots\}$ by the following procedure. Let $v_1 \in U$ be arbitrary. If $v_1, \ldots, v_m$ are already defined then consider the balls $B_k(v_i)$ of radius k around the v_i, i.e., $B_k(v_i)$ consists of those elements of U that differ from v_i in at most k coordinates. If

$$U_m := U \setminus \bigcup_{i=1}^{m} B_k(v_i) \neq \emptyset$$

then we define v_{m+1} as an arbitrary element of U_m; if $U_m = \emptyset$ then we stop the construction of V.

Using (3.7), we can estimate $|B_k(v_i)|$ as

$$|B_k(v_i)| = \sum_{0 \le j \le k} \binom{5k}{j} < (k+1)\frac{(5k)!}{k!(4k)!} < 3(k+1)\sqrt{5k}\left(\frac{5^5}{4^4}\right)^k < 2^{4k-1}$$

where the last inequality holds for n larger than an absolute constant. Therefore, if $m \le 2^{k+1}$ then $U_m \neq \emptyset$, proving that $|V| > 2^{k+1} > n$. □

The following lemma is the main step toward answering the question raised at the beginning of the subsection.

Lemma 4.5. *Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$, $e \in A$ and $\langle A \rangle = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. Then there are $g \in A^n$, $h \in A^{\lfloor n^{44 \log n} \rfloor}$ such that the action of $\langle g, h \rangle$ on $[n]$ has at most $472(\log n)^2$ orbits, provided that n is larger than an absolute constant.*

Proof. Let $k = 5\lfloor (\log n)/(\log 2) \rfloor$. By Lemma 4.3 there is an element $g \in A^n$ with $|\mathrm{supp}(g)| = \alpha n \ge n/2$. Let $\varepsilon = 1/n$ and $\ell = \lceil 2n^{6k} \log(n^{2k}/\varepsilon) \rceil$. Let $h \in A^{\ell}$ be the outcome of a random walk of length ℓ as in Lemma 4.2.

Consider all words of the form

$$f(\vec{a}) = hg^{a_1} hg^{a_2} \ldots hg^{a_k},$$

where $\vec{a} = (a_i : 1 \le i \le k)$ runs through all sequences in $U = \{0,1\}^k$. For $\beta \in [n]$, we wish to estimate from below the length of the orbit $\beta^{\langle g,h \rangle}$ by counting the number of different images $f_\beta(\vec{a}) := \beta^{f(\vec{a})}$, for $\vec{a} \in U$.

To this end, for fixed elements $\vec{a} = (a_1, \ldots, a_k)$ and $\overrightarrow{a'} = (a'_1, \ldots, a'_k)$ in U and $\beta \in [n]$, we wish to bound from above the probability that $f_\beta(\vec{a}) = f_\beta(\overrightarrow{a'})$. We will do this by examining all possible trajectories $(\beta_1, \ldots, \beta_k)$, $(\beta'_1, \ldots, \beta'_k)$, where

$$\beta_1 = \beta^{hg^{a_1}},\ \beta_2 = \beta_1^{hg^{a_2}},\ \ldots,\ \beta_k = \beta_{k-1}^{hg^{a_k}} \quad \text{and} \quad \beta'_1 = \beta^{hg^{a'_1}},\ \ldots,\ \beta'_k = \beta'^{hg^{a'_k}}_{k-1},$$

counting how many satisfy $\beta_k \neq \beta'_k$, and then estimating the probability (for h chosen randomly in the manner described above) that such a pair of trajectories be traversed following $f(\vec{a})$ and $f(\vec{a'})$.

Let $R = \{1 \leq i \leq k : a_i \neq a_i'\}$; let the elements of R be $k_1 < k_2 < \cdots < k_r$, where $r = |R|$. Let $r_0 \leq r$ be fixed. Let $k' = k_{r_0}$. Consider all tuples $(\beta_1, \beta_2, \ldots, \beta_k, \beta_{k'}', \ldots, \beta_k') \in [n]^{(2k-k')+1}$ such that

(a) $\beta_1, \beta_2, \ldots, \beta_k, \beta_{k'}', \ldots, \beta_k'$ are distinct from each other and from β,

(b) $\beta_1^{g^{-a_1}}, \beta_2^{g^{-a_2}}, \ldots, \beta_k^{g^{-a_k}}$, $(\beta_{k'+1}')^{g^{-a_{k'+1}'}}, \ldots, (\beta_k')^{g^{-a_k'}}$ are distinct from each other,

(c) $\beta_{k_j} \notin \mathrm{supp}(g)$ for every $j < r_0$, but $\beta_{k'} \in \mathrm{supp}(g)$,

(d) $(\beta_{k'}')^{g^{-a_{k'}'}} = \beta_k^{g^{-a_{k'}}}$.

The number of such tuples is at least

$$\left(\prod_{j=1}^{r_0-1} (n - |\mathrm{supp}(g)| - j)\right) \cdot (|\mathrm{supp}(g)| - 1) \cdot \prod_{j=(r_0+1)}^{2k-k'} (n - (2j-1)), \tag{4.1}$$

where we count tuples by choosing first $\beta_{k_j} \in [n] \setminus \mathrm{supp}(g)$ for $1 \leq j < r_0$, then $\beta_{k'} \in \mathrm{supp}(g)$, then the other β_i and β_i'. To justify the estimate on the number of choices at each stage, notice that at the j^{th} choice with $j \leq r_0 - 1$ we have to make selections from $[n] \setminus \mathrm{supp}(g)$ so as to satisfy (c) while keeping them different from previous selections and from β (to satisfy (a)). Then $\beta_{k'}$ can be chosen as an arbitrary element of $\mathrm{supp}(g)$ different from β. At this point, (b) is still satisfied automatically. At later choices, if β_i or β_i' is selected at stage j then enforcing (a) eliminates j possibilities and enforcing (b) eliminates $j-1$, not necessarily different, possibilities. Note that (4.1) also gives a valid lower estimate (namely, 0) in the case when $r_0 - 1 \geq n - |\mathrm{supp}(g)|$.

By Lemma 4.2 (with $2k$ instead of k, and with properties (a), (b) as inputs), the probability that a random $h \in A^\ell$ satisfies

$$(\beta, \beta_1, \ldots, \beta_{k-1}, \beta_{k'}', \ldots, \beta_{k-1}')^h = (\beta_1^{g^{-a_1}}, \beta_2^{g^{-a_2}}, \beta_k^{g^{-a_k}}, (\beta_{k'+1}')^{g^{-a_{k'+1}'}}, \ldots, (\beta_k')^{g^{-a_k'}}) \tag{4.2}$$

is at least $(1-\varepsilon)\frac{(n-(2k-k'))!}{n!} > (1-\varepsilon)\frac{1}{n^{2k-k'}}$. If h satisfies (4.2) then $\beta^{hg^{a_1}} = \beta_1$, $\beta_1^{hg^{a_2}} = \beta_2, \ldots, \beta_{k-1}^{hg^{a_k}} = \beta_k$. By properties (c) and (d), we also have $\beta^{hg^{a_1'}} = \beta_1$, $\beta_1^{hg^{a_2'}} = \beta_2, \ldots, \beta_{k'-1}^{hg^{a_{k'}'}} = \beta_{k'}^{g^{a_{k'}'-a_{k'}}} = \beta_{k'}'$; by (4.2), we also have $\left(\beta_{k'}'\right)^{hg^{a_{k'+1}'}} = \beta_{k'+1}'$, $\ldots$, $\left(\beta_{k-1}'\right)^{hg^{a_k'}} = \beta_k'$. Thus, in particular, any two distinct tuples

$$(\beta_1, \beta_2, \ldots, \beta_k, \beta_{k'}', \ldots, \beta_k')$$

give us mutually exclusive events, even for different values of r_0. Note also that, by property (a) and what we have just said, $f_\beta(\vec{a}) = \beta_k \neq \beta_k' = f_\beta(\vec{a'})$.

Hence the probability P that $f_\beta(\vec{a}) \neq f_\beta(\vec{a'})$ is at least

$$\begin{aligned} P &\geq \sum_{r_0=1}^{r} \frac{1-\varepsilon}{n^{2k-k_{r_0}}} \cdot \left(\prod_{j=1}^{r_0-1}(n - \alpha n - j)\right) \cdot (\alpha n - 1) \cdot \prod_{j=(r_0+1)}^{2k-k_{r_0}} (n - (2j-1)) \\ &> \sum_{r_0=1}^{r} \left(1 - \tfrac{1}{n}\right) \left(1 - \tfrac{4k}{n}\right)^{2k} \left(\alpha - \tfrac{1}{n}\right) \cdot \prod_{j=1}^{r_0-1} \left(1 - \alpha - \tfrac{j}{n}\right). \end{aligned} \tag{4.3}$$

If $\alpha n = |\operatorname{supp}(g)| \geq n-k$ then we estimate P from below by the summand $r_0 = 1$ in (4.3), yielding

$$P > \left(1-\frac{1}{n}\right)\left(1-\frac{4k}{n}\right)^{2k}\left(\alpha-\frac{1}{n}\right) > 1-\frac{1}{n}-\frac{8k^2}{n}-\frac{k+1}{n} > 1-\frac{9k^2}{n},$$

with the last inequality valid if n is larger than an absolute constant.

If $\alpha n = |\operatorname{supp}(g)| < n-k$ then, estimating the terms $(1-\alpha-j/n)$ in the last product in (4.3) from below by $(1-\alpha-k/n)$, we obtain

$$\begin{aligned} P &> \left(1-\frac{1}{n}\right)\left(1-\frac{4k}{n}\right)^{2k}\left(\alpha-\frac{1}{n}\right)\sum_{r_0=1}^{r}\left(1-\alpha-\frac{k}{n}\right)^{r_0-1} \\ &> \left(1-\frac{4k}{n}\right)^{2k+1}\left(\alpha-\frac{1}{n}\right)\frac{1-(1-\alpha-(k/n))^r}{(1-(1-\alpha-(k/n))} \qquad (4.4) \\ &= \left(1-\frac{4k}{n}\right)^{2k+1}\frac{\alpha-(1/n)}{\alpha+(k/n)}\left(1-(1-\alpha-(k/n))^r\right). \end{aligned}$$

Since $\alpha \geq 1/2$, we have $\frac{\alpha-(1/n)}{\alpha+(k/n)} \geq 1-\frac{2(k+1)}{n}$ and $(1-\alpha-(k/n))^r < (1/2)^r$, implying

$$P > 1-\frac{4k(2k+1)}{n}-\frac{2(k+1)}{n}-\frac{1}{2^r} > 1-\frac{9k^2}{n}-\frac{1}{2^r},$$

if n is larger than an appropriate absolute constant.

We conclude that, for any two non-identical tuples

$$\vec{a} = (a_1,\ldots,a_k) \in \{0,1\}^k, \quad \vec{a'} = (a'_1,\ldots,a'_k) \in \{0,1\}^k$$

and for any $\beta \in [n]$,

$$\operatorname{Prob}(\beta^{hg^{a_1}hg^{a_2}\ldots hg^{a_k}} = \beta^{hg^{a'_1}hg^{a'_2}\ldots hg^{a'_k}}) < \frac{9k^2}{n}+\frac{1}{2^{r(\vec{a},\vec{a'})}},$$

where $r(\vec{a},\vec{a'})$ is the number of indices $1 \leq j \leq k$ for which $a_j \neq a'_j$.

By Lemma 4.4, there exists a set V of more than n tuples so that any two tuples differ in more than $k/5$ coordinates. For fixed $\beta \in [n]$, writing $f_\beta(\vec{a}) = \beta^{hg^{a_1}hg^{a_2}\ldots hg^{a_k}}$, $\vec{a} \in V$ for the random variable $\beta \mapsto f_\beta(\vec{a})$ defined using a random $h \in A^\ell$, we obtain that

$$\begin{aligned} &\mathbb{E}(|\{(\vec{a},\vec{a}') \in V^2 : f_\beta(\vec{a}) = f_\beta(\vec{a}')\}|) = \sum_{\vec{a},\vec{a}' \in V}\operatorname{Prob}(f_\beta(\vec{a}) = f_\beta(\vec{a}')) \\ &\leq |V| + \left(\frac{9k^2}{n}+\frac{1}{2^{r(\vec{a},\vec{a'})}}\right)|V|(|V|-1) < \frac{|V|^2}{n}+\left(\frac{9k^2}{n}+\frac{1}{2^{k/5}}\right)|V|^2 \\ &< (9k^2+3)\frac{|V|^2}{n} < 472(\log n)^2\frac{|V|^2}{n}. \end{aligned}$$

Concerning the length of the orbit $\beta^{\langle g,h\rangle}$, we have

$$\begin{aligned}\mathbb{E}\left(\frac{1}{|\beta^{\langle g,h\rangle}|}\right) &\leq \mathbb{E}\left(\frac{1}{|\{f_\beta(\vec{a}) : \vec{a} \in V\}|}\right)\\ &\leq \mathbb{E}\left(\frac{|\{(\vec{a},\vec{a}') \in V^2 : f_\beta(\vec{a}) = f_\beta(\vec{a}')\}|}{|V|^2}\right) \leq \frac{472(\log n)^2}{n},\end{aligned}$$

where we use the Cauchy-Schwartz inequality (or the inequality between the arithmetic and quadratic means) in the second step for the numbers m_i that measure how many times a particular value γ_i occurs among the $f_\beta(\vec{a})$, for some $\vec{a} \in V$.

Now, $\sum_{\beta\in[n]} 1/|\beta^{\langle g,h\rangle}|$ is just the number of orbits of $\langle g,h\rangle$ (since each such orbit contributes $|\beta^{\langle g,h\rangle}| \cdot 1/|\beta^{\langle g,h\rangle}| = 1$ to the sum). Hence

$$\mathbb{E}(\text{number of orbits of } \langle g,h\rangle) \leq 472(\log n)^2.$$

In particular, there exists an $h \in A^\ell$ such that the number of orbits of $\langle g,h\rangle$ is at most $472(\log n)^2$. Since $\ell = \lceil 2n^{6k}\log(n^{2k}/\varepsilon)\rceil$ is smaller than $n^{44\log n}$ for n larger than a constant, we are done. □

Proposition 4.6. *Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$, $e \in A$ and $\langle A\rangle = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. If n is larger than an absolute constant, then there are $g_1, g_2, g_3 \in A^{\lfloor n^{44\log n}\rfloor}$ such that $\langle g_1, g_2, g_3\rangle$ is transitive.*

Proof. Let g, h be as in Lemma 4.5. Let $\varepsilon = 1/n^2$, $\ell = \lceil 2n^6\log(n^2/\varepsilon)\rceil$. Let $g' \in A^\ell$ be the outcome of a random walk of length ℓ as in Lemma 4.2. Note that $\ell \leq \lfloor n^{44\log n}\rfloor$ for n larger than an absolute constant.

Let Δ be the union of orbits of $\langle g,h\rangle$ of length less than $\sqrt{n}$. Since, by Lemma 4.5, there are at most $472(\log n)^2$ orbits of $\langle g,h\rangle$, we have $|\Delta| < 472\sqrt{n}(\log n)^2$. Let S be a set consisting of one element α of each orbit of length less than $\sqrt{n}$. Then, for each $\alpha \in S$, Lemma 4.2 implies that

$$\mathrm{Prob}\left(\alpha^{g'} \in \Delta\right) \leq (1+\varepsilon)\frac{|\Delta|}{n} < \left(1 + \frac{1}{n^2}\right)\frac{472(\log n)^2}{\sqrt{n}}$$

and so

$$\mathrm{Prob}\left((\exists\alpha \in S)\ (\alpha^{g'} \in \Delta)\right) < \left(1 + \frac{1}{n^2}\right)\frac{472^2(\log n)^4}{\sqrt{n}}. \tag{4.5}$$

Let κ be an orbit of $\langle g,h\rangle$ contained in $n \setminus \Delta$; by definition, $|\kappa| \geq \sqrt{n}$. Let κ_0 be the largest orbit; by the pigeonhole principle, $|\kappa_0| > \frac{n}{472(\log n)^2}$. Then

$$\mathbb{E}(|\kappa^{g'} \cap \kappa_0|) = \sum_{\alpha\in\kappa}\mathrm{Prob}(\alpha^{g'} \in \kappa_0) \geq \sum_{\alpha\in\kappa}(1-\varepsilon)\frac{|\kappa_0|}{n} = (1-\varepsilon)\frac{|\kappa||\kappa_0|}{n},$$

whereas

$$
\begin{aligned}
\mathbb{E}\left(|\kappa^{g'} \cap \kappa_0|^2\right) &= \sum_{\alpha,\beta\in\kappa} \text{Prob}\left(\alpha^{g'} \in \kappa_0 \wedge \beta^{g'} \in \kappa_0\right) \\
&= \sum_{\alpha\in\kappa} \text{Prob}(\alpha^{g'} \in \kappa_0) + \sum_{\substack{\alpha,\beta\in\kappa \\ \alpha\neq\beta}} \sum_{\substack{\alpha',\beta'\in\kappa_0 \\ \alpha'\neq\beta'}} \text{Prob}((\alpha,\beta)^{g'} = (\alpha',\beta')) \\
&\leq \sum_{\alpha\in\kappa} (1+\varepsilon)\frac{|\kappa_0|}{n} + \sum_{\alpha,\beta\in\kappa,\alpha\neq\beta} (1+\varepsilon)\frac{|\kappa_0|(|\kappa_0|-1)}{n(n-1)} \\
&\leq (1+\varepsilon)\left(\frac{|\kappa_0||\kappa|}{n} + \frac{|\kappa|(|\kappa|-1)|\kappa_0|(|\kappa_0|-1)}{n(n-1)}\right) \\
&\leq (1+\varepsilon)\left(\frac{|\kappa_0||\kappa|}{n} + \frac{|\kappa|^2|\kappa_0|^2}{n^2}\right).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\text{Var}(|\kappa^{g'} \cap \kappa_0|) &= \mathbb{E}(|\kappa^{g'} \cap \kappa_0|^2) - \mathbb{E}(|\kappa^{g'} \cap \kappa_0|)^2 \\
&\leq (1+\varepsilon)\left(\frac{|\kappa_0||\kappa|}{n} + \frac{|\kappa_0|^2|\kappa|^2}{n^2}\right) - (1-\varepsilon)^2\frac{|\kappa_0|^2|\kappa|^2}{n^2} \\
&\leq 3\varepsilon\frac{|\kappa|^2|\kappa_0|^2}{n^2} + (1+\varepsilon)\frac{|\kappa_0||\kappa|}{n} < \left(1+\frac{4}{n}\right)\frac{|\kappa_0||\kappa|}{n}.
\end{aligned}
$$

By Chebyshev's inequality,

$$
\begin{aligned}
\text{Prob}(\kappa^{g'} \cap \kappa_0 = \emptyset) &\leq \frac{\text{Var}(|\kappa^{g'} \cap \kappa_0|)}{\mathbb{E}(|\kappa^{g'} \cap \kappa_0|)^2} \\
&\leq \frac{(|\kappa||\kappa_0|/n)(1+4/n)}{(1-\varepsilon)^2\frac{|\kappa|^2|\kappa_0|^2}{n^2}} \leq \frac{12n}{|\kappa||\kappa_0|} < \frac{12\cdot 472(\log n)^2}{\sqrt{n}}.
\end{aligned}
$$

Hence

$$
\text{Prob}\left((\exists\kappa \subseteq ([n]\setminus\Delta))\ (\kappa^{g'} \cap \kappa_0 = \emptyset)\right) < \frac{12\cdot 472^2(\log n)^4}{\sqrt{n}}. \tag{4.6}
$$

Now, for n larger than a constant,

$$
\left(1+\frac{1}{n^2}\right)\frac{472^2(\log n)^4}{\sqrt{n}} + \frac{12\cdot 472^2(\log n)^4}{\sqrt{n}} < 1.
$$

Therefore, (4.5) and (4.6) imply that with positive probability, (a) $\kappa^{g'}$ intersects $[n]\setminus\Delta$ for every orbit κ not contained in $[n]\setminus\Delta$ and (b) $\kappa^{g'}$ intersects κ_0 for every orbit κ contained in $[n]\setminus\Delta$. In particular, this happens for some $g' \in A^\ell$. Properties (a) and (b) imply that $\langle g, h, g'\rangle$ is transitive. We set $g_1 = g$, $g_2 = h$, $g_3 = g'$ and are done. ∎

We will later use[7] the following corollary with $k = 2$.

Corollary 4.7. *Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$, $e \in A$ and $\langle A\rangle = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. Let $k \geq 1$. If n is larger than a constant depending only on k, then there is a set $S \subseteq A^{\lfloor n^{45 \log n}\rfloor}$ of size at most $3k$ such that $\langle S\rangle$ is k-transitive.*

Proof. Let $\alpha_1 \in [n]$ be arbitrary. Since $\langle A\rangle$ is transitive, Lemma 3.9 implies that $\alpha_1^{A^n} = [n]$. Let $G = \mathrm{Sym}([n])$, $H = G_{\alpha_1}$, $A' = A^n$. Since $\alpha_1^{A'} = [n]$, A' intersects every coset of H in G. By Schreier's Lemma (Lem 3.8), it follows that $(A')^3 \cap H$ generates $\langle A\rangle \cap H$, which is either $\mathrm{Sym}([n] \setminus \{\alpha_1\})$ or $\mathrm{Alt}([n] \setminus \{\alpha_1\})$. Let $A_1 = (A')^3 \cap H$.

Iterating, we obtain a sequence of sets $A_0 = A, A_1, A_2, \ldots, A_{k-1} \subseteq \mathrm{Sym}([n])$ and a sequence of elements $\alpha_1, \alpha_2, \ldots, \alpha_{k-1} \in [n]$ such that $A_i \subseteq A_{i-1}^{3n}$ and $\langle A_i\rangle$ is $\mathrm{Sym}([n] \setminus \{\alpha_1, \ldots, \alpha_i\})$ or $\mathrm{Alt}([n] \setminus \{\alpha_1, \ldots, \alpha_i\})$.

Let $(g_1)_i$, $(g_2)_i$, $(g_3)_i$ be as in Prop. 4.6, applied with A_i instead of A. Then $(g_1)_i, (g_2)_i, (g_3)_i \in A^{(3n)^i \lfloor n^{44 \log n}\rfloor}$ and $\langle (g_1)_i, (g_2)_i, (g_3)_i\rangle \subseteq \mathrm{Sym}([n] \setminus \{\alpha_1, \ldots, \alpha_i\})$ is transitive on $[n] \setminus \{\alpha_1, \ldots, \alpha_i\}$ for $0 \leq i \leq k-1$. Thus, for $S = \bigcup_{i=0}^{k-1} A_i$, $\langle S\rangle$ is k-transitive on $[n]$. □

5. The splitting lemma and its consequences

We will prove what is in effect an adaptation of Babai's splitting lemma (proven for groups in [Bab82, Lem. 3.1]) to the case of sets. This is a key point in this paper: the splitting lemma will allow us to construct long stabiliser chains with large orbits.

The following easy lemma will make an "unfolding" step possible.

Lemma 5.1. *Let $A \subseteq \mathrm{Sym}([n])$, $\Sigma \subseteq [n]$ and $g \in \mathrm{Sym}([n])$. Then*

$$gA_{(\Sigma^g)}g^{-1} = (gAg^{-1})_{(\Sigma)}.$$

Proof. We have $\mathrm{Sym}([n])_{(\Sigma^g)} = g^{-1}\,\mathrm{Sym}([n])_{(\Sigma)}g$. Therefore,

$$\begin{aligned} A_{(\Sigma^g)} &= A \cap \mathrm{Sym}([n])_{(\Sigma^g)} = A \cap g^{-1}\,\mathrm{Sym}([n])_{(\Sigma)}g \\ &= g^{-1}(gAg^{-1} \cap \mathrm{Sym}([n])_{(\Sigma)})g = g^{-1}(gAg^{-1})_{(\Sigma)}g. \end{aligned}$$

□

Notice a feature of the following statement – there is a high power of A in the assumptions, not just in the conclusion. We will "unfold" the high power of A in the course of the proof. (By Σ^S we mean the set $\Sigma^S = \{\alpha^g : \alpha \in \Sigma, g \in S\}$.)

Proposition 5.2 (Splitting Lemma)**.** *Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$, $e \in A$ and $\langle A\rangle$ 2-transitive. Let $\Sigma \subseteq [n]$. Assume that there are at least $\rho n(n-1)$ ordered pairs (α, β) of distinct elements of $[n]$ such that there is no $g \in (A^{\lfloor 5n^6 \log n\rfloor})_{(\Sigma)}$ with $\alpha^g = \beta$. Then there is a subset S of $A^{\lfloor 3n^6 \log n\rfloor}$ with*

$$(AA^{-1})_{(\Sigma^S)} = \{e\}$$

[7]If we wished to, we could use it to obtain a set S of generators of $\mathrm{Alt}([n])$ or $\mathrm{Sym}([n])$ simply by setting $k = 6$: the Classification of Finite Simple Groups implies that a 6-transitive group must be either alternating or symmetric.

and

$$|S| \leq \left\lceil \frac{2}{\log(3/(3-2\rho))} \cdot \log n \right\rceil.$$

Proof. Set $\ell = \lceil 2n^6 \log(n^2/(1/3)) \rceil$; note that $\ell \leq \lfloor 3n^6 \log n \rfloor$ and $2\ell+2 \leq \lfloor 5n^6 \log n \rfloor$ for $n \geq 5$. (For $n < 5$, the statement is trivial.) By Lemma 4.2 applied with $k = 2$ and $\varepsilon = 1/3$, we obtain that given any two distinct elements $\alpha, \beta \in [n]$ and $g \in A^\ell$, the pair (α^g, β^g) adopts any possible value (α', β') with probability at least $(1 - 1/3)/(n(n-1))$, where we choose $g \in A^\ell$ with the distribution in Lemma 4.2 ($g = g_1 g_2 \cdots g_\ell$, g_i chosen independently from $A' \cup \{e\}$, where A' is a symmetric subset of A). Since this distribution is symmetric, this is the same as saying that $(\alpha^{g^{-1}}, \beta^{g^{-1}})$ adopts any possible value (α', β') with probability at least $(1 - 1/3)/(n(n-1))$.

Now, given (α, β) and $g \in A^\ell$, we have $h \in (AA^{-1})_{(\Sigma^g)}$ and $\alpha^h = \beta$ if and only if $ghg^{-1} \in g(AA^{-1})_{(\Sigma^g)} g^{-1}$ and $(\alpha^{g^{-1}})^{ghg^{-1}} = \beta^{g^{-1}}$. By Lemma 5.1 applied to AA^{-1}, we have that $ghg^{-1} \in g(AA^{-1})_{(\Sigma^g)} g^{-1}$ only if $ghg^{-1} \in (gAA^{-1}g^{-1})_{(\Sigma)}$, which in turn can happen only if $ghg^{-1} \in (A^{2\ell+2})_{(\Sigma)}$. Thus, if there is no element $j \in (A^{2\ell+2})_{(\Sigma)}$ with $\alpha^{g^{-1}j} = \beta^{g^{-1}}$, then there is no element $h \in (AA^{-1})_{(\Sigma^g)}$ with $\alpha^h = \beta$. (This is the "unfolding" step we referred to before.)

Since by hypothesis there are at least $\rho n(n-1)$ ordered pairs (α', β') such that there is no element $j \in (A^{2\ell+2})_{(\Sigma)}$ with $\alpha'^j = \beta'$, and since $(\alpha^{g^{-1}}, \beta^{g^{-1}})$ equals any such pair with probability at least $(2/3)/(n(n-1))$, we see that the probability that there is no element $h \in (AA^{-1})_{(\Sigma^g)}$ with $\alpha^h = \beta$ is at least $2\rho/3$.

Let S be a set of r random $g \in A^\ell$ (chosen independently, with the distribution as above). The probability that for every $g \in S$ there is an element $h \in (AA^{-1})_{(\Sigma^g)}$ with $\alpha^h = \beta$ is at most $(1 - 2\rho/3)^r$. This must happen if there is an element $h \in (AA^{-1})_{\Sigma^S}$ such that $\alpha^h = \beta$. Thus, the probability that there is such an h is at most $(1-2\rho/3)^r$, and the probability that there is such an h for at least one of the $n(n-1)$ pairs (α, β) is at most $n(n-1)(1 - 2\rho/3)^r$.

Setting $r = \lceil (\log n^2)/(\log 3/(3 - 2\rho)) \rceil$, we obtain that the probability that there is such an h for at least one pair is less than 1. Hence there is a set $S \subseteq A^\ell$ with at most r elements such that, for every pair (α, β) of distinct elements of $[n]$, there is no $h \in (AA^{-1})_{(\Sigma^S)}$ with $\alpha^h = \beta$. This implies immediately that the only element of $(AA^{-1})_{(\Sigma^S)}$ is the identity. □

Corollary 5.3. *Let $A \subseteq \mathrm{Sym}([n])$ with $A = A^{-1}$, $e \in A$ and $\langle A \rangle$ 2-transitive. Let $A' = A^{\lfloor 5n^6 \log n \rfloor}$. Let $\Sigma \subseteq [n]$ be such that*

$$|\alpha^{A'_{(\Sigma)}}| < (1 - \rho)n$$

for every $\alpha \in [n]$, where $\rho \in (0, 1)$. Then

$$|\Sigma| > \frac{\log |A|}{\left\lceil \frac{2}{\log(3/(3-2\rho))} \cdot \log n \right\rceil \cdot \log n}.$$

In particular, if $\rho = 0.05$ then $|\Sigma| > (\log |A|)/(60(\log n)^2)$.

Proof. Since $|\alpha^{A'_{(\Sigma)}}| < (1-\rho)n$ for every $\alpha \in [n]$, there are at least $\rho n(n-1)$ tuples (α,β) such that there is no $g \in A'$ with $\alpha^g = \beta$. By Prop. 5.2, there is a set $S \subseteq \mathrm{Sym}([n])$ such that $(AA^{-1})_{(\Sigma^S)} = \{e\}$ and $|S| \leq \left\lceil \frac{2}{\log(3/(3-2\rho))} \cdot \log n \right\rceil$. Since $(AA^{-1})_{(\Sigma^S)} = \{e\}$, we know, by Lemma 3.17, that $|\Sigma^S| \geq \log_n |A|$. Clearly $|\Sigma^S| \leq |S||\Sigma|$. Hence

$$|\Sigma| \geq \frac{\log_n |A|}{|S|} \geq \frac{\log |A|}{\left\lceil \frac{2}{\log(3/(3-2\rho))} \cdot \log n \right\rceil \cdot \log n}.$$

□

A key idea in the proof of the Main Theorem is the following. For $A \subseteq \mathrm{Sym}([n])$, we can construct $A' = A^{\lfloor 5n^6 \log n \rfloor}$ and a set $\Sigma = \{\alpha_1, \alpha_2, \dots\} \subseteq [n]$ starting with an empty set and taking at each step α_i to be an element such that $|\alpha_i^{(A')_{(\alpha_1,\dots,\alpha_{i-1})}}| \geq (1-\rho)n$ (say); if no such element exists, we stop the procedure. By Cor. 5.3, $|\Sigma|$ must be large.

An application of Lemma 3.19 will give that, for $A'' = (A')^{16n^6}$, the set $(A'')_\Sigma$ contains a copy of $\mathrm{Alt}(\Delta)$, where $\Delta \subseteq \Sigma$ and $|\Delta| \geq (1-\rho)|\Sigma|$. Such a large alternating group certainly looks like a valuable tool.

6. Proof of the main theorem

The core of this section is Proposition 6.5. It is a growth result, but not quite of type $|A \cdot A \cdot A| \geq |A|^{1+\varepsilon}$ or $|A^k| \geq |A|^{1+\varepsilon}$. What will grow by a factor at each step is not the number of elements $|A|$ of A, but rather the length m of a sequence $\alpha_1, \dots, \alpha_m$ such that the orbits

$$\alpha_1^A, \alpha_2^{A_{\alpha_1}}, \alpha_3^{A_{(\alpha_1,\alpha_2)}}, \dots, \alpha_m^{A_{(\alpha_1,\alpha_2,\dots,\alpha_{m-1})}} \tag{6.1}$$

are all large.

This growth result (Prop. 6.5) will be applied iteratively. There are two ways for the iteration to stop: (a) an element we construct could fix a large set pointwise (we call this the case of *exit*), or (b) a group we work with could fail to have a large alternating composition factor. In case (a), we obtain all of $G = \mathrm{Alt}([n])$ in a few steps by Thm. 1.4. In case (b), we can *descend* to the problem of proving small diameter for n' smaller than n by a constant factor. (Here, as is "infinite descent", the term "descent" means the same as induction, seen backwards.)

* * *

Let us sketch briefly the proof of Prop. 6.5. First, we use (6.5) to construct many elements in the setwise stabiliser G_Σ, where $\Sigma = \{\alpha_1, \dots, \alpha_m\}$; in fact we get an entire copy of a large alternating group in $(G_\Sigma)|_\Sigma$ (Lemma 3.19). This is the *setup*. Then comes the *creation* step: we use the action by conjugation of G_Σ on the pointwise stabiliser $G_{(\Sigma)}$ to construct many elements of $G_{(\Sigma)}$ (Lemma 6.1). We *organise* these new elements (all in a power A' of A) as follows: we apply Cor. 5.3 (a consequence of the splitting lemma) to lengthen our stabiliser chain $A' \supseteq A'_{\alpha_1} \supseteq \dots \supseteq A'_{(\alpha_1,\dots,\alpha_m)} \supseteq \dots$ up to $A'_{(\alpha_1,\dots,\alpha_{m+\ell})}$ in such a way that the orbits (defined as

in (6.1)) are still large. We repeat the *organiser* step about $\gg (\log n)/(\log m)$ times. There are only two ways for this procedure to stop prematurely, namely, *exit* and *descent* (cases (a) and (b) discussed above).

* * *

We start by proving the lemma containing the *creation* step: we give a way to construct many elements in a subgroup H^- of a group G. The basic idea is the application of the orbit-stabiliser principle to the action by conjugation of a subgroup $H^+ \leq N_G(H^-)$ on H^-, where $N_G(H^-)$ is the normaliser of H^-.

Lemma 6.1. *Let $G = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$, $H^- \leq G$, $H^+ \leq N_G(H^-)$, Γ an orbit of H^+. Let $Y = \{y_1, \dots, y_r\} \subseteq H^-$ be such that $\langle Y\rangle|_\Gamma$ is 2-transitive on Γ. Let $B \subseteq H^+$. Then either*

(a) *there is a $b \in BB^{-1} \setminus \{e\}$ fixing Γ pointwise, or*
(b) $|B^{-1}YB \cap H^-| \geq |B|^{1/r}$.

Proof. Consider the action of B on $\vec{y} = (y_1, \dots, y_r)$ by conjugation: for $b \in B$, we define $\vec{y}^b := (y_1^b, \dots, y_r^b)$, where $y^b = b^{-1}yb$. Assume first that there are two distinct elements $b_1, b_2 \in B$ such that $\vec{y}^{b_1}|_\Gamma = \vec{y}^{b_2}|_\Gamma$. Then $b_1b_2^{-1}|_\Gamma$ centralises $\vec{y}|_\Gamma$, implying that $b_1b_2^{-1}|_\Gamma \in Z(\langle Y\rangle|_\Gamma) = \{e\}$. Hence $b_1b_2^{-1} \in B$ fixes Γ pointwise without being the identity, i.e., conclusion (a) holds.

Assume now that the restrictions $\vec{y}^b|_\Gamma$ are all distinct. Hence, by the pigeonhole principle, there exists an index $j \in \{1, \dots, r\}$ such that the set W of conjugates of y_j by B satisfies $|W|_\Gamma| \geq |B|^{1/r}$. Observe that all elements of W are in H^-, as $Y \subset H^-$ and $B \subset N(H^-)$. Hence $|B^{-1}YB \cap H^-| \geq |W| \geq |B|^{1/r}$. □

The following useful lemma is in part an easy application of Schreier's lemma and in part a consequence of a trick based on the following trivial fact: one clearly cannot have two disjoint copies within $[n]$ of an orbit of size greater than $n/2$.

Lemma 6.2. *Let $\Delta \subseteq [n]$. Let $B^+ \subseteq (\mathrm{Sym}(n))_\Delta$ with $B^+ = (B^+)^{-1}$, $e \in B^+$. Assume $B^+|_\Delta$ is $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$. Let $B^- = \left((B^+)^3\right)_{(\Delta)}$.*

Then $\langle B^-\rangle = \langle B^+\rangle_{(\Delta)} \lhd \langle B^+\rangle$. Furthermore, if $\langle B^-\rangle$ has an orbit Γ of length greater than $n/2$, then Γ is also an orbit of $\langle B^+\rangle$.

Proof. Since $B^+|_\Delta$ is a group ($\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$), $B^+|_\Delta = \langle B^+\rangle|_\Delta$. Thus B^+ contains an element from every coset of $\langle B^+\rangle_{(\Delta)}$ in $\langle B^+\rangle$ and so, by Lemma 3.8, B^- contains a set of generators of $\langle B^+\rangle_{(\Delta)}$. Hence $\langle B^-\rangle = \langle B^+\rangle_{(\Delta)}$. In particular, $\langle B^-\rangle \lhd \langle B^+\rangle$, as $\langle B^+\rangle_{(\Delta)}$ is the kernel of the action of $\langle B^+\rangle$ on Δ.

The orbits of the normal subgroup $\langle B^-\rangle \lhd \langle B^+\rangle$ are blocks of imprimitivity for $\langle B^+\rangle$. Since one cannot have two blocks of length greater than $n/2$, $\langle B^+\rangle$ leaves Γ invariant as a set, and so Γ is an orbit of $\langle B^+\rangle$. □

Our *descent* step rests in part on a simple calculation.

Lemma 6.3. *Let $f(x) = e^{(\log x)^k g(x)}$, where $k \in \mathbb{Z}^+$, $t > 0$, and $g : [t, \infty) \to [0, \infty)$ is an increasing function. Let $0 < \delta_0 < 1$ and $\delta_1, \delta_2 > 0$ be given with $\delta_1 + \delta_2 <$*

$k\log(1/\delta_0)$. Let $h:[t,\infty)\to[0,\infty)$ be a function with $h(x)\leq e^{\delta_2(\log x)^{k-1}g(x)}$. Then

$$f(\delta_0 x)\cdot h(x)\leq f(x)^{1-\frac{\delta_1}{\log x}} \tag{6.2}$$

provided x is larger than a bound depending only on k, t, δ_0, δ_1, and δ_2.

Proof. Since the derivative of the function $j(z):=(1-z\log(1/\delta_0))^k$ at $z=0$ is less than $-(\delta_1+\delta_2)$, there exists an interval $(0,\varepsilon)$, with ε depending on k, δ_0, δ_1, and δ_2, such that $j(z)<1-(\delta_1+\delta_2)z$ holds for all $z\in(0,\varepsilon)$. Hence, if x is large enough that $1/\log x<\varepsilon$ then

$$(\log(\delta_0 x))^k=(\log x)^k\cdot j\left(\frac{1}{\log x}\right)<(\log x)^k\left(1-\frac{\delta_1+\delta_2}{\log x}\right).$$

If in addition $\delta_0 x\geq t$ so that $f(\delta_0 x)$, $f(x)$, $g(x)$, and $h(x)$ are defined then

$$\begin{aligned}f(\delta_0 x)&=e^{(\log\delta_0 x)^k g(\delta_0 x)}\leq e^{(\log\delta_0 x)^k g(x)}\leq e^{((\log x)^k-(\delta_1+\delta_2)(\log x)^{k-1})g(x)}\\&\leq(h(x))^{-1}e^{-\delta_1(\log x)^{k-1}g(x)}\cdot f(x)=(h(x))^{-1}f(x)^{1-\frac{\delta_1}{\log x}}.\end{aligned}$$

□

The following lemma is also crucial to the descent step. In the proof of the lemma, we use Lemma 3.20 to guarantee the existence of an element that we then construct by other means.

Lemma 6.4. *Let $G=\mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. Let $\Delta\subseteq[n]$, $|\Delta|\geq(\log n)^2$. Let $A\subseteq G$ with $A=A^{-1}$, $e\in A$ and $\langle A\rangle=G$. Let $B^+\subseteq(A^l)_\Delta$, $l\geq 1$, with $B^+=(B^+)^{-1}$, $e\in B^+$. Assume $B^+|_\Delta$ is $\mathrm{Alt}(\Delta)$ or $\mathrm{Sym}(\Delta)$. Let $B^-=\left((B^+)^3\right)_{(\Delta)}$. Assume $\langle B^-\rangle$ has an orbit Γ of length at least ρn, for some $\rho>8/9$.*

If all alternating composition factors $\mathrm{Alt}(k)$ of $\langle B^-\rangle$ satisfy $k\leq\delta n$, where $\delta>0$, and

$$\max_{k\leq\delta n}\mathrm{diam}(\mathrm{Alt}(k))\leq D_\delta, \tag{6.3}$$

for some $D_\delta>0$, and n is larger than an absolute constant, then

$$A^{\lfloor \ell e^{c(\log n)^3}\cdot D_\delta\rfloor}\supseteq\mathrm{Alt}([n]),$$

where $c=c(\rho)$ depends only on ρ.

Proof. The group $U:=\langle B^-\rangle|_\Gamma$ is transitive. It is also isomorphic to a factor group of $\langle B^-\rangle$, so U also has no alternating factors $\mathrm{Alt}(k)$ with $k>\delta n$. By Thm. 1.1 and by (6.3), there exists an absolute constant C_1 such that for

$$u:=\lfloor e^{C_1(\log n)^3}\cdot D_\delta\rfloor,\quad (B^-)^u|_\Gamma=U. \tag{6.4}$$

Let $H=\langle B^+\rangle$. By Lemma 6.2, Γ is an orbit of H. If n is large enough that Lemma 3.20 applies then there exists $g\in H$ of support less than $|\Gamma|/4$ on Γ. Take $h\in B^+$ with $h|_\Delta=g|_\Delta$. Then $gh^{-1}\in\langle B^+\rangle_{(\Delta)}=\langle B^-\rangle$ and so, by (6.4), there exists $b\in(B^-)^u$ with $gh^{-1}|_\Gamma=b|_\Gamma$. Therefore, $bh\in(B^+)^{3u+1}$ satisfies $bh|_\Gamma=g|_\Gamma$. Since g fixes at least $(3/4)|\Gamma|\geq(3/4)\cdot\rho n>(2/3)n$ points in Γ, we have $|\mathrm{supp}(bh)|\leq(1-(3/4)\rho)n<n/3$. By Thm. 1.4, $(A\cup\{bh,(bh)^{-1}\})^{Kn^8}$ contains

Alt($[n]$), where $K = K(\varepsilon)$ ($\varepsilon = 1-(3/4)\rho < 1/3$) is the number defined in Thm. 1.4. Since $A \cup \{bh, (bh)^{-1}\} \subseteq A^{(3u+1)l}$, we are done. $\square$

We come to the key result in this section.

Proposition 6.5. *Let $G = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. Let $A \subset G$ with $A = A^{-1}$, $e \in A$, and $\langle A\rangle = G$. Let $\alpha_1, \alpha_2, \ldots, \alpha_{m+1} \in [n]$ be such that*

$$\left|\alpha_i^{A_{(\alpha_1,\ldots,\alpha_{i-1})}}\right| \geq \frac{9}{10} n \tag{6.5}$$

for every $i = 1, 2, \ldots, m+1$.

There exist absolute constants $C > 1$, $c_1, c_2, c_3, c_4 > 0$ with

$$c_4/c_3 < c_2 < 4(\log 1/0.95)c_1 \tag{6.6}$$

such that the following holds: if $m \geq (\log n)^2$, then either

$$A^{\left\lfloor Ce^{c_1(\log n)^4 \log\log n - c_2(\log n)^3 \log\log n}\right\rfloor} \supseteq \mathrm{Alt}([n]) \tag{6.7}$$

or there are $\alpha_{m+2}, \alpha_{m+3}, \ldots, \alpha_{m+\ell+1} \in [n]$, $\ell \geq c_3(m \log m)/(\log n)$, such that

$$\left|\alpha_i^{A'_{(\alpha_1,\ldots,\alpha_{i-1})}}\right| \geq \frac{9}{10} n \tag{6.8}$$

for $A' = A^{\lfloor n^{c_4 \log n}\rfloor}$ and every $i = 1, 2, \ldots, m+\ell+1$.

The condition $c_4/c_3 < c_2 < 4(\log 1/0.95)c_1$ is needed for the recursion to work. As we will see, this inequality is easy to attain; we are not taking advantage of any numerical coincidence. We may choose $c_1 = 50182$, $c_2 = 10296$, $c_3 = 0.0745$, and $c_4 = 767$.

An easy application of Proposition 6.5 proves Corollary 6.6 (which is equivalent to our Main Theorem). Conversely, in order to prove Proposition 6.5, we will use Corollary 6.6 for smaller values of n in an inductive process. In the proofs of Prop. 6.5 and Cor. 6.6, we assume that n is greater than a well-defined (but not explicitly computed) absolute constant n_0; we need that n is greater than an appropriate constant so n satisfies some inequalities we introduce during the proofs, and also so that results from the previous sections, valid only for sufficiently large n, can be applied. Once n_0 is defined, the constant C in the statement of Prop. 6.5 can be adjusted so that (6.7) holds trivially for $n \leq n_0$.

Corollary 6.6. *Let $G = \mathrm{Sym}([n])$ or $\mathrm{Alt}([n])$. Let $Y \subseteq G$ with $Y = Y^{-1}$, $e \in Y$ and $G = \langle Y\rangle$. Then $\mathrm{diam}(\Gamma(G,Y)) \leq Ce^{c_1(\log n)^4 \log\log n}$, where C and c_1 are as in Proposition 6.5.*

The proof consists just of a repeated use of Proposition 6.5, plus some accounting.

Proof. We can assume that n is large enough that $m_0 \leq 0.1n \leq n-3$ for $m_0 = \lfloor(\log n)^2\rfloor + 1$, and so G acts transitively on the set X of all (m_0+1)-tuples. Hence, by Lemma 3.9, the set $A_0 := Y^{n^{m_0+1}} \supseteq Y^{|X|}$ acts transitively on the set of all (m_0+1)-tuples. Thus (6.5) holds with A_0 instead of A, m_0 instead of m and $\alpha_i = i$

for $i = 1, 2, \ldots, m_0+1$. We apply Proposition 6.5 with these parameters. We obtain either (6.7) or (6.8).

In the latter case, we set $\ell_0 = \ell$, $m_1 = m_0 + \ell_0$, and iterate: we apply Proposition 6.5 to

$$A_1 = A_0^r, A_2 = A_1^r = A_0^{r^2}, A_3 = A_2^r = A_0^{r^3}, \ldots$$

where $r = \lfloor n^{c_4 \log n} \rfloor$. (After each step, we "save" the output ℓ to ℓ_i and set $m_{i+1} = m_i + \ell_i$.) We stop when we obtain (6.7); say this happens when we apply Proposition 6.5 with $A = A_k = A_0^{r^k}$. Clearly $m_k \leq 0.1n$ (as otherwise (6.5) could not hold for $i = m_k + 1$).

It remains to estimate k. For $1 \leq j \leq \log n$, let t_j be the largest index i between 0 and k such that $m_i < e^j$; if no such index exists, set $t_j = 1$. We have $m_0 \geq 3$ and so $t_1 = 1$. By Proposition 6.5, $m_{i+1} \geq (1 + (c_3 \log m_i)/(\log n)) \cdot m_i$ for all $i < k$, and so, by $(1 + c_3 j/(\log n))^{\lfloor (\log n)/(c_3 j)\rfloor + 2} > e$, we have $t_{j+1} \leq t_j + \lfloor (\log n)/(c_3 j)\rfloor + 2$. Thus

$$t_{\lfloor \log n \rfloor} \leq t_1 + \sum_{j=1}^{\lfloor \log n\rfloor - 1} (t_{j+1} - t_j) \leq 1 + \sum_{1 \leq j \leq \log n} \left(\frac{\log n}{c_3 j} + 2\right) \leq c_5 \log n \log\log n$$

for any $c_5 > 1/c_3$, with the last inequality valid if n is larger than an appropriate constant. Set $c_5 = (1/c_3 + c_2/c_4)/2$ (say); note that (6.6) ensures that $c_2/c_4 > c_3$, and so $c_2/c_4 > c_5 > 1/c_3$. Since $t_{\lfloor \log n\rfloor} + 1 > k$ (because $m_k \leq 0.1n$), we get that $k \leq c_5 \log n \log\log n$.

Thus

$$A_k = A_0^{r^k} \subseteq Y^{n^{\lfloor (\log n)^2\rfloor + 2} \cdot r^{\lfloor c_5 \log n \log\log n\rfloor}} \subseteq Y^{\lfloor e^{c_6 (\log n)^3 \log\log n}\rfloor},$$

for any $c_6 > c_4 \cdot c_5$, if n is large enough in terms of c_6. Set $c_6 = (c_4 \cdot c_5 + c_2)/2$ (say), so that $c_6 < c_2$. Then, by (6.7) (valid for $A = A_k$), we obtain

$$\begin{aligned} \mathrm{Alt}([n]) &\subseteq (Y^{\lfloor e^{c_6 (\log n)^3 \log\log n}\rfloor})^{\lfloor C e^{c_1 (\log n)^4 \log\log n - c_2 (\log n)^3 \log\log n}\rfloor} \\ &\subseteq Y^{\lfloor C e^{c_1 (\log n)^4 \log\log n}\rfloor - 1} \end{aligned}$$

for n larger than a constant. If $Y \subseteq \mathrm{Alt}([n])$, then $Y^{\lfloor C e^{c_1 (\log n)^4 \log\log n}\rfloor - 1} = \mathrm{Alt}([n])$. If Y contains an odd permutation then $Y^{\lfloor C e^{c_1 (\log n)^4 \log\log n}\rfloor} = \mathrm{Sym}([n])$. □

We finally turn to the proof of Proposition 6.5. In what follows, let

$$f(x) = e^{c_1 (\log x)^4 \log\log x} \quad \textit{and} \quad g(x) = c_1 \log\log x. \tag{6.9}$$

Proof of Proposition 6.5. We prove Proposition 6.5 by induction on n. During the proof, in the inductive hypothesis, we will make use of Cor. 6.6 for values smaller than n.

We can assume that n is large enough so that $m \geq (\log n)^2 > C(0.9)$, where $C(0.9)$ is as in Lemma 3.19. Apply Lemma 3.19 with $d = 0.9$ and $\Sigma = \{\alpha_1, \ldots, \alpha_m\}$. We obtain a set $\Delta \subseteq \Sigma$ such that $|\Delta| \geq 0.9|\Sigma|$ and $\left(\left(A^{16m^6}\right)_\Sigma\right)_{(\Sigma\setminus\Delta)}|_\Delta$ contains $\mathrm{Alt}(\Delta)$.

Let

$$B^+ = \left\{ g \in \left(\left(A^{16m^6} \right)_\Sigma \right)_{(\Sigma \setminus \Delta)} : g|_\Delta \in \mathrm{Alt}(\Delta) \right\}, \quad B^- = \left((B^+)^3 \right)_{(\Delta)}.$$

This is our initial *setup*: we have a large set B^+ in the setwise stabiliser G_Σ; furthermore, we have constructed a large subset $\Delta \subseteq \Sigma$ such that $B^+ \subseteq (G_\Sigma)_{(\Sigma\setminus\Delta)}$ and $B^+|_\Delta = \mathrm{Alt}(\Delta)$. We also have a set B^- in the pointwise stabiliser $G_{(\Sigma)}$. By (6.5) with $i = m+1$, $\left|\alpha_{m+1}^{B^-}\right| \geq \frac{9}{10}n$, and so $\langle B^- \rangle$ has an orbit Γ of length at least $0.9n$. By Lemma 6.2, Γ is also an orbit of $\langle B^+ \rangle$.

We would like $\langle B^- \rangle$ to act as an alternating or symmetric group on Γ; let us show that, if this is not the case, we obtain *descent*. We may assume that Corollary 6.6 holds for $n' < n$ (inductive hypothesis). Hence, if $\langle B^- \rangle$ has no composition factor $\mathrm{Alt}(k)$ with $k > 0.95n$, then Lemma 6.4 (*descent*) gives us

$$A^{\lfloor 16m^6 e^{c(\log n)^3} \cdot Cf(0.95n) \rfloor} \supseteq \mathrm{Alt}([n]),$$

for n larger than an absolute constant, where $c = c(0.9)$ is from Lemma 6.4 and $f(x)$ is as in (6.9). We apply Lemma 6.3 with $k = 4$, $h(x) = 16x^6 e^{c(\log x)^3}$, $\delta_0 = 0.95$, $\delta_1 = c_2/c_1$, $\delta_2 \in (0, 4\log(1/\delta_0) - \delta_1)$ arbitrary, and t large enough that $h(x) \leq \exp(\delta_2 c_1 (\log x)^3 \log\log x)$ holds for $x \geq t$. We obtain that $16m^6 e^{c(\log n)^3} \cdot Cf(0.95n) \leq Cf(n)^{1-(c_2/c_1)/(\log x)}$ for n larger than a constant, and so (6.7) holds and we are done.

Thus, we can suppose from now on that $\langle B^- \rangle$ does have a composition factor $\mathrm{Alt}(k)$ for some $k > 0.95n$. The only orbit of $\langle B^- \rangle$ that can be of length at least k is Γ, so $\langle B^- \rangle|_\Gamma = \langle B^-|_\Gamma \rangle$ must contain $\mathrm{Alt}(k)$ as a section. Hence, by Lemma 3.11, $\langle B^-|_\Gamma \rangle \geq \mathrm{Alt}(\Gamma)$. (We can assume $0.95n > 84$, and thus Lemma 3.11 does apply.) Note we also get that $|\Gamma| > 0.95n$.

Now that we know that $\langle B^-|_\Gamma \rangle \geq \mathrm{Alt}(\Gamma)$, Corollary 4.7 gives us a small set of elements $Y = \{y_1, y_2, \dots, y_6\} \subseteq (B^-)^{\lfloor n^{45 \log n} \rfloor}$ such that $\langle Y \rangle|_\Gamma$ is 2-transitive on Γ. We apply Lemma 6.1 (*creation*) with $H^- = \langle B^- \rangle$, $H^+ = \langle B^+ \rangle$, $B = B^+$ and $r = 6$. (The condition $H^- \lhd H^+$ is fulfilled thanks to Lemma 6.2.)

If conclusion (a) in Lemma 6.1 holds, then there is a $b \in B^+(B^+)^{-1} \setminus \{e\}$ with $\mathrm{supp}(b) \leq 0.05n$. Thm. 1.4 thus gives us that $(A \cup \{b\})^{Kn^8} \supseteq \mathrm{Alt}([n])$, where $K = K(0.05)$ is an absolute constant. Assuming that n is large enough to satisfy

$$2 \cdot 48m^6 \cdot Kn^8 < 96Kn^{14} < \lfloor \exp\left(c_1 (\log n)^4 \log\log n - c_2 (\log n)^3 \log\log n\right) \rfloor,$$

we obtain that (6.7) holds and we are done. (This is what we call an *exit* from the procedure.)

We can thus assume that conclusion (b) in Lemma 6.1 holds, i.e., we have *created* a set $W = (B^+)^{-1} Y B^+ \cap \langle B^- \rangle$ with $|W| \geq |B^+|^{1/6}$. Note that $(B^+)^{-1} Y B^+ \subset A^{\lfloor n^{46 \log n} \rfloor}$ (for n larger than a constant) and $|B^+| \geq |\mathrm{Alt}(\Delta)| = (1/2)|\Delta|! \geq m^{0.899m}$ (for m larger than a constant; recall that $|\Delta| \geq 0.9m$). Hence

$$\left| A^{\lfloor n^{46 \log n} \rfloor} \cap \langle B^- \rangle \right| \geq m^{0.149m}. \tag{6.10}$$

Now that we have *created* many elements in the pointwise stabiliser of Σ, it is our task to *organise* them: we wish to produce $\alpha_{m+2}, \dots, \alpha_{m+\ell+1}$ satisfying (6.8).

For $i \geq 0$, we define recursively $A_i, B_i \subseteq \langle A \rangle$ and a sequence Σ_i of points in $[n]$. Let $A_0 = A^{\lfloor n^{46 \log n} \rfloor}$, $m_0 = m$, $\Sigma_0 = (\alpha_1, \ldots, \alpha_{m_0+1})$, and $B_0 = (A_0)_{(\Sigma_0 \setminus \{\alpha_{m_0+1}\})}$.

If A_i, Σ_i, B_i are already defined then let $A'_{i+1} = A_i^{\lfloor 5n^6 \log n \rfloor}$ and let Σ_{i+1} be a largest possible extension $\Sigma_{i+1} = (\alpha_1, \ldots, \alpha_{m_{i+1}+1})$ of $\Sigma_i = (\alpha_1, \ldots, \alpha_{m_i+1})$ such that

$$\left| \alpha_j^{(A'_{i+1})_{(\alpha_1,\ldots,\alpha_{j-1})}} \right| \geq 0.9n, \tag{6.11}$$

for all $j = 1, 2, \ldots, m_{i+1} + 1$. Finally, let

$$A_{i+1} = (A'_{i+1})^{48n^6} \quad \text{and} \quad B_{i+1} = (A_{i+1})_{(\Sigma_{i+1} \setminus \{\alpha_{m_{i+1}+1}\})}.$$

Note that for all $i \geq 0$, $\langle B_i \rangle$ has an orbit Γ_i of length at least $0.9n$ because $\left|\alpha_{m_i+1}^{B_i}\right| \geq 0.9n$.

We stop the recursion, and set $w := i$ for the last i for which A_i is defined, if either

(a) $|B_i|_{\Gamma_i}| < |B_i|$, i.e., there are two elements $b_1, b_2 \in B_i$ such that $b_1 b_2^{-1}$ fixes Γ_i pointwise; or
(b) $|\Gamma_i| \leq 0.95n$ or $\langle B_i|_{\Gamma_i} \rangle \not\supset \mathrm{Alt}(\Gamma_i)$ or
(c) $n^{m_i - m_0} > \sqrt{m^{0.149m}}$.

By (6.10), we have $|B_0| \geq m^{0.149m}$.

First, we estimate the differences $m_{i+1} - m_i$. If the recursion did not stop after the definition of A_i, B_i, and Σ_i then, in particular, the stopping criterion (c) is not fulfilled at step i. Lemma 3.4, applied with $\langle B_0 \rangle$ as G, $G_{(\Sigma_i \setminus \Sigma_0)}$ as H, and B_0 as A, then implies that

$$|B_i| \geq |B_0^2 \cap H| \geq \frac{|B_0|}{n^{m_i - m_0}} \geq \sqrt{m^{0.149m}}.$$

Also, by the criteria (a) and (b), we have $|B_i|_{\Gamma_i}| = |B_i|$ and $\langle B_i \rangle|_{\Gamma_i}$ acts as $\mathrm{Alt}(\Gamma_i)$ or $\mathrm{Sym}(\Gamma_i)$ on Γ_i, where $|\Gamma_i| > 0.95n$.

Since $0.9n < 0.95 \cdot 0.95n \leq 0.95|\Gamma_i|$, we can apply Corollary 5.3 with $\rho = 0.05$, $B_i|_{\Gamma_i}$ instead of A, and Γ_i instead of $[n]$, and obtain that, for $1 \leq i < w$,

$$m_{i+1} - m_i > \frac{\log |B_i|}{60(\log n)^2} \geq \frac{c_3 m \log m}{60(\log n)^2}, \tag{6.12}$$

where we define $c_3 := 0.149/2 = 0.0745$. (This is what we have called an *organiser* step. It is ultimately based on the splitting lemma (Prop. 5.2), of which Cor. 5.3 is a corollary.)

At the same time, $n^{m_{w-1} - m_0} \leq \sqrt{m^{0.149m}}$ implies

$$m_{w-1} - m_0 \leq \frac{c_3 \log m}{\log n} m.$$

Since $m_{w-1} - m_0 = \sum_{i=1}^{w-1} (m_i - m_{i-1})$, from (6.12) it follows that

$$\frac{c_3 \log m}{\log n} m > (w-1) \frac{c_3 m \log m}{60(\log n)^2}$$

and we conclude that $w-1 < 60\log n$. Hence

$$A_w = A_0^{\lfloor 5n^6 \log n\rfloor^w (48n^6)^w} \subseteq A^{\lfloor n^{46\log n}\rfloor \cdot \lfloor 240 n^{12}\log n\rfloor^w} \subseteq A^{\lfloor n^{c_4 \log n}\rfloor}$$

for $c_4 := 767 > 46 + 12\cdot 60$, provided that n is larger than an absolute constant.

If $n^{m_w - m_0} > \sqrt{m^{0.149m}}$ (stopping condition (c)), then

$$m_w - m_0 \geq \frac{c_3 \log m}{\log n} m,$$

and so, setting $\ell = m_w - m_0$, we obtain (6.8). (In other words, our repeated *organising* has succeeded.)

If we stopped because condition (a) holds then A_w^2 contains a non-trivial element $b_1 b_2^{-1}$ with support less than $0.1n$. By Theorem 1.4, $(A \cup \{b_1 b_2^{-1}\})^{Kn^8} \supseteq \mathrm{Alt}([n])$, where $K = K(0.1)$ is an absolute constant. Assuming that n is large enough to satisfy

$$2\cdot \lfloor n^{c_4 \log n}\rfloor \cdot Kn^8 < \left\lfloor e^{c_1 (\log n)^4 \log\log n - c_2 (\log n)^3 \log\log n}\right\rfloor,$$

we obtain (6.7). (This is an *exit* case.)

Finally, suppose we stopped in case (b), i.e., $\langle B_w|_{\Gamma_w}\rangle \not\supseteq \mathrm{Alt}(\Gamma_w)$ or $|\Gamma_w| \leq 0.95n$. As $|\Sigma_w| \geq m > C(0.9)$, we can apply Lemma 3.19 with Σ_w as Σ and A'_w as A, to obtain $\Delta_w \subseteq \Sigma_w$, $|\Delta_w| \geq 0.9|\Sigma_w|$ and $B_w^+ \subseteq (A'_w)^{16n^6}$ such that B_w^+ fixes $\Sigma_w \setminus \Delta_w$ pointwise, fixes Δ_w setwise, and $B^+|_{\Delta_w} = \mathrm{Alt}(\Delta_w)$. (This is a fresh *setup*.) Also, by Lemma 6.2, $B_w^- = \left((B_w^+)^3\right)_{(\Delta_w)}$ generates $\langle B_w^+\rangle_{(\Delta_w)} \lhd \langle B_w^+\rangle$. Note that $B_w^- \subseteq B_w$ and $\langle B_w^-\rangle$ has an orbit of length at least $0.9n$, because $B_w^- \supseteq \alpha_{m_w+1}^{A'_w}$.

We are ready for another *descent*. The group $\langle B_w^-\rangle$ has no composition factor $\mathrm{Alt}(k)$ with $k > 0.95n$, because such a factor would be a section of $\langle B_w\rangle$ and Lemma 3.11 would imply that $\langle B_w|_{\Gamma_w}\rangle$ is an alternating group on $> 0.95n$ elements, in contradiction with condition (b). Thus the hypotheses of Lemma 6.4 are satisfied with $\delta = 0.95$ and $\rho = 0.9$ and, by the inductive hypothesis, Lemma 6.4 gives us that

$$A^{\lfloor n^{c_4 \log n} e^{c(\log n)^3}\cdot Cf(0.95n)\rfloor} \supseteq \mathrm{Alt}([n]),$$

where $c = c(0.9)$ and f is as in (6.9). We apply Lemma 6.3 with $k = 4$, $h(x) = x^{c_4 \log x} e^{c(\log x)^3}$, $\delta_0 = 0.95$, $\delta_1 = c_2/c_1$, $\delta_2 \in (0, 4\log(1/\delta_0) - \delta_1)$ arbitrary, and t large enough that $h(x) \leq \exp(\delta_2 c_1 (\log x)^3 \log\log x)$ holds for $x \geq t$. We obtain that $n^{c_4 \log n} e^{c(\log n)^3}\cdot Cf(0.95n) \leq Cf(n)^{1-(c_2/c_1)/(\log x)}$ for n larger than a constant, and so we obtain conclusion (6.7).

We end by noting that we can set $c_2 = \lceil c_4/c_3\rceil = 10296$, $c_1 = \lceil c_2/(4\log 1/0.95)\rceil = 50182$, and thus fulfill (6.6). $\square$

We now use Corollary 6.6 to prove both the Main Theorem and Cor. 1.3 (for $\mathrm{Sym}(n)$ and $\mathrm{Alt}(n)$) with explicit constants in the exponents.

Theorem 6.7. *Let $G = \mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$. Then*

$$\mathrm{diam}(G) = O(e^{c_1 (\log n)^4 \log\log n}),$$
$$\overrightarrow{\mathrm{diam}}(G) = O(e^{(c_1+1)(\log n)^4 \log\log n}),$$

where $c_1 > 0$ is the absolute constant in Corollary 6.6.

As we have said, $c_1 = 50182$ is valid.

Proof. Let A be an arbitrary set of generators of G. Let $X = A \cup A^{-1} \cup \{e\}$. The undirected Cayley graph $\Gamma(G, X)$ is just the undirected Cayley graph $\Gamma(G, A)$ with a loop at every vertex; their diameters are the same. Thus, by Corollary 6.6,

$$\mathrm{diam}(\Gamma(G, A)) = \mathrm{diam}(\Gamma(G, X)) \leq Ce^{c_1 (\log n)^4 \log \log n}.$$

By [Bab06, Cor. 2.3],

$$\mathrm{diam}(\vec{\Gamma}(G, A)) \leq O\left(\mathrm{diam}(G)(n \log n)^2\right) \leq O\left(e^{(c_1+1)(\log n)^4 \log \log n}\right).$$

□

Finally, we note that, as $Ce^{c_1 (\log n)^4 \log \log n} \leq e^{(c_1 + \log C)(\log n)^4 \log \log n}$, the bounds obtained in Thm. 6.7 can be written in the form stated in the Main Theorem and Cor. 1.3.

References

[Ald87] D. Aldous. On the Markov chain simulation method for uniform combinatorial distributions and simulated annealing. *Prob. Engng. Info. Sci.*, 1(1):33–46, 1987.

[Bab91] L. Babai. Local expansion of vertex transitive graphs and random generation in finite groups. In *23rd ACM Symposium on Theory of Computing*, pages 164–174. ACM Press, New York, NY, 1991.

[Bab06] L. Babai. On the diameter of Eulerian orientations of graphs. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 822–831, New York, 2006. ACM.

[Bab82] L. Babai. On the order of doubly transitive permutation groups. *Invent. Math.*, 65(3):473–484, 1981/82.

[BBS04] L. Babai, R. Beals, and Á. Seress. On the diameter of the symmetric group: polynomial bounds. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1108–1112 (electronic), New York, 2004. ACM.

[BG08a] J. Bourgain and A. Gamburd. On the spectral cap for finitely generated subgroups of SU(2). *Invent. Math.*, 171:83–121, 2008.

[BG08b] J. Bourgain and A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[BGH+] J. Bamberg, N. Gill, T. Hayes, H. Helfgott, G. Royle, Á. Seress, and P. Spiga. Bounds on the diameter of Cayley graphs of the symmetric group. Preprint.

[BGS] J. Bourgain, A. Gamburd, and P. Sarnak. Generalization of Selberg's 3/16 theorem and affine sieve. Available as `arxiv.org:0912.5021`.

[BGS10] J. Bourgain, A. Gamburd, and P. Sarnak. Affine linear sieve, expanders, and sum-product. *Invent. Math.*, 179(3):559–644, 2010.

[BGT] E. Breuillard, B. Green, and T. Tao. Approximate subgroups of linear groups. To appear in *GAFA*. Available as `arxiv.org:1005.1881`.

[BH05] L. Babai and T. Hayes. Near-independence of permutations and an almost sure polynomial bound on the diameter of the symmetric group. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1057–1066. ACM, New York, 2005.

[BHK+90] L. Babai, G. Hetyei, W. M. Kantor, A. Lubotzky, and Á. Seress. On the diameter of finite groups. In *31st Annual Symposium on Foundations of Computer Science, Vol. I,*

II (St. Louis, MO, 1990), pages 857–865. IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.

[Boc89] A. Bochert. Über die Zahl der verschiedenen Werthe, die eine Funktion gegenüber Buchstaben durch Vertauschung derselben erlangen kann. *Math. Ann.*, 33:584–590, 1889.

[BS87] L. Babai and Á. Seress. On the degree of transitivity of permutation groups: a short proof. *J. Combin. Theory Ser. A*, 45(2):310–315, 1987.

[BS88] L. Babai and Á. Seress. On the diameter of Cayley graphs of the symmetric group. *J. Combin. Theory Ser. A*, 49(1):175–179, 1988.

[BS92] L. Babai and Á. Seress. On the diameter of permutation groups. *European J. Combin.*, 13(4):231–243, 1992.

[Din] O. Dinai. Growth in SL_2 over finite fields. To appear in *J. Group Theory.*

[DM96] J. D. Dixon and B. Mortimer. *Permutation groups*, volume 163 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[DSC93] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.*, 21(4):2131–2156, 1993.

[Fie72] M. Fiedler. Bounds for eigenvalues of doubly stochastic matrices. *Linear Algebra and Appl.*, 5:299–310, 1972.

[Gan91] A. Gangolli. *Convergence bounds for Markov chains and applications to sampling.* PhD thesis, Dept. Computer Science, Stanford Univ., 1991.

[GHa] N. Gill and H. Helfgott. Growth in small generating sets in $\mathrm{SL}_n(\mathbb{Z}/p\mathbb{Z})$. To appear in *IMRN*. Available as `arxiv.org:1002.1605`.

[GHb] N. Gill and H. Helfgott. Growth in solvable subgroups of $\mathrm{GL}_r(\mathbb{Z}/p\mathbb{Z})$. Submitted. Available as `arxiv.org:1008.5264`.

[Hel08] H. A. Helfgott. Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. of Math. (2)*, 167(2):601–623, 2008.

[Hel11] H. A. Helfgott. Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$. *J. Eur. Math. Soc. (JEMS)*, 13(3):761–851, 2011.

[Hru] E. Hrushovski. Stable group theory and approximate subgroups. Available as `arxiv.org:0909.2190`.

[KMS84] D. Kornhauser, G. Miller, and P. Spirakis. Coordinating pebble motion on graphs, the diameter of permutation groups, and applications. In *Proceedings of the 25th IEEE Symposium on Foundations of Computer Science*, pages 241–250, Singer Island, FL, 1984. IEEE Computer Society Press.

[Lie83] M. W. Liebeck. On graphs whose full automorphism group is an alternating group or a finite classical group. *Proc. London Math. Soc. (3)*, 47(2):337–362, 1983.

[Lov96] L. Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 353–397. János Bolyai Math. Soc., Budapest, 1996.

[LP11] M. J. Larsen and R. Pink. Finite subgroups of algebraic groups. *J. Amer. Math. Soc.*, 24(4):1105–1158, 2011.

[McK84] P. McKenzie. Permutations of bounded degree generate groups of polynomial diameter. *Inform. Process. Lett.*, 19(5):253–254, 1984.

[Moh91] B. Mohar. Eigenvalues, diameter, and mean distance in graphs. *Graphs Combin.*, 7(1):53–64, 1991.

[Pak] I. Pak. Problems: new, old and unusual. Talk at National University of Ireland, Galway, Ireland, Dec 1, 2009. Available as `http://larmor.nuigalway.ie/~detinko/Igor.pdf`.

[PPSS] C. Praeger, L. Pyber, P. Spiga, and E. Szabó. The Weiss conjecture for locally primitive graphs with automorphism groups admitting composition factors of bounded rank. To appear in *Proc. Amer. Math. Soc.*

[PS] L. Pyber and E. Szabó. Growth in finite simple groups of Lie type of bounded rank. Submitted. Available as `arxiv.org:1005.1881`.

[PS80] Ch. E. Praeger and J. Saxl. On the orders of primitive permutation groups. *Bull. London Math. Soc.*, 12(4):303–307, 1980.

[Pyb93] L. Pyber. On the orders of doubly transitive permutation groups, elementary estimates. *J. Combin. Theory Ser. A*, 62(2):361–366, 1993.

[RT85] I. Z. Ruzsa and S. Turjányi. A note on additive bases of integers. *Publ. Math. Debrecen*, 32(1-2):101–104, 1985.

[Ser03] Á. Seress. *Permutation Group Algorithms*, volume 152 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2003.

[SGV] A. Salehi-Golsefidy and P. Varjú. Expansion in perfect groups. Submitted. Available as `arxiv.org:1108.4900`.

[Sim70] Ch. C. Sims. Computational methods in the study of permutation groups. In *Computational Problems in Abstract Algebra (Proc. Conf., Oxford, 1967)*, pages 169–183. Pergamon, Oxford, 1970.

[SSV05] B. Sudakov, E. Szemerédi, and V. H. Vu. On a question of Erdős and Moser. *Duke Math. J.*, 129(1):129–155, 2005.

[Tao08] T. Tao. Product set estimates for non-commutative groups. *Combinatorica*, 28(5):547–594, 2008.

[Var] P. Varjú. Expansion in $\mathrm{SL}_d(\mathcal{O}_K/I)$, I square-free. Submitted. Available as `arxiv.org:1001.3664`.

Harald A. Helfgott, École Normale Supérieure, Département de Mathématiques, 45 rue d'Ulm, F-75230 Paris, France
E-mail address: `harald.helfgott@ens.fr`

Ákos Seress, Centre for the Mathematics of Symmetry and Computation, The University of Western Australia, Crawley, WA 6009 Australia, and Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA
E-mail address: `akos@math.ohio-state.edu`